

1542
(4-85)

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE

ONLINE SEARCH REQUEST FORM

8-393

 USER Sheela Huff SERIAL NUMBER 08/23/85
 ART UNIT 1806 PHONE 30578666 DATE 8/26/96

Please give a detailed statement of requirements. Describe as specifically as possible the subject matter to be searched. Define any terms that may have special meaning. Give examples or relevant citations, authors, or keywords, if known.

You may include a copy of the broadest and or relevant claim(s).

Please search Seq. 10 # 18, 2
and 26, 27

TXS-

Sheela

Also Reverse transcribe AA seq's and search
on NA databases.

STAFF USE ONLY

COMPLETED 8/22/96
 SEARCHER mqk
 ONLINE TIME 6 TOTAL TIME 10
(in minutes)
 NO. OF DATABASES 6

SYSTEMS
☐ CAS ONLINE
☐ DARC/QUESTEL
☐ DIALOG
☐ SDC
☒ OTHER MPSACTT

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch nn n.a. - n.a. database search, using Smith-Waterman algorithm

```
Run on: Tue Aug 27 08:32:33 1996; MasPar time 625.55 Seconds
896.015 Million cell updates/sec
```

Tabular output not generated.

Title: >US-08-231-565A-1

Description: (1-1559) from US08231565A.seq

Perfect Score: 1559
N.A. Sequence: 1 AGCAGACAGAGGACTCTCAT.....AAAAAAAAAAAAAAAAAAAA 1559

TCGTCGTCTCCTGAGAGTA.....TTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0: Querv 0

Searched: 514334 seqs. 179763086 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: EST-STs

1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97
98:EST98 99:EST99 100:EST100

Database:

98: EST91 99: EST92 100: EST93 101: EST94 102: EST95
103: EST96 104: EST97 105: EST98 106: EST99 107: EST100
108: EST101 109: gNEST7 110: gNEST2 111: gNEST3 112: gNEST4
113: gNEST5 114: gNEST6 115: gNEST7 116: gNEST8 117: gNEST9
118: gNEST10 119: gNEST11 120: gNEST12 121: gNEST13
122: gNEST14 123: gNEST15 124: gNEST1 125: gNEST2 126: gNEST3
127: gNEST2 128: gNEST3 129: gNEST4 130: gNEST5 131: gNEST6

Statistics:
Mean 11.935: Variance 4.909: scale 2.431

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description	Pred. No.	
	No.	Score	Match				
c	1	519	33.3	642 134	HS587267	yx91e07.s1 Homo sapie	0.00e+00
c	2	519	33.3	642 114	N26587	yx91e07.s1 Homo sapie	0.00e+00
c	3	488	31.3	593 113	N23770	yx35a12.s1 Homo sapie	0.00e+00
c	4	488	31.3	593 136	HS770262	yx35a12.s1 Homo sapie	0.00e+00
c	5	482	30.9	522 134	HS510255	yw11h07.s1 Homo sapie	0.00e+00
c	6	482	30.9	522 111	HS7510	yw11h07.s1 Homo sapie	0.00e+00
c	7	468	30.0	540 111	HS7498	yw11f07.s1 Homo sapie	0.00e+00
c	8	468	30.0	540 133	HS498253	yw11f07.s1 Homo sapie	0.00e+00
c	9	441	28.3	473 112	N20091	yx36a12.s1 Homo sapie	0.00e+00
c	10	441	28.3	473 129	HS091255	yx36a12.s1 Homo sapie	0.00e+00
c	11	423	27.1	433 113	N24677	yx91a09.s1 Homo sapie	0.00e+00
c	12	423	27.1	433 135	HS677260	yx91a09.s1 Homo sapie	0.00e+00
c	13	394	25.3	396 136	HS796259	yx97f10.s1 Homo sapie	0.00e+00
c	14	394	25.3	396 113	N24796	yx97f10.s1 Homo sapie	0.00e+00
c	15	392	25.1	401 114	N26562	yx62h05.s1 Homo sapie	0.00e+00
c	16	392	25.1	401 134	HS562267	yx62h05.s1 Homo sapie	0.00e+00
c	17	357	22.9	365 118	N39656	yx91e07.s1 Homo sapie	0.00e+00
c	18	357	22.9	365 135	HS656272	yx91e07.s1 Homo sapie	0.00e+00
c	19	357	22.9	370 117	N35680	yx62h05.s1 Homo sapie	0.00e+00
c	20	357	22.9	370 135	HS680270	yx62h05.s1 Homo sapie	0.00e+00
c	21	353	22.6	393 135	HS656271	yx91a09.s1 Homo sapie	0.00e+00
c	22	353	22.6	393 117	N36656	yx91a09.s1 Homo sapie	0.00e+00
c	23	289	18.5	303 82	T27884	EST19370 Homo sapiens	1.28e-249
c	24	224	14.4	406 131	HS298280	yy23g10.s1 Homo sapie	4.46e-184
c	25	223	14.3	226 130	HS198272	yy25h10.s1 Homo sapie	4.48e-183
c	26	223	14.3	226 116	N32198	yy25h10.s1 Homo sapie	4.48e-183
c	27	217	13.9	349 79	T01292	seq1461 Homo sapiens	4.52e-177
c	28	216	13.9	505 23	H64579	yy63h07.s1 Homo sapie	4.52e-176
c	29	215	13.8	370 44	M77888	EST01472 Homo sapiens	4.52e-175
c	30	215	13.8	387 44	M77964	EST01548 Homo sapiens	4.52e-175
c	31	215	13.8	445 130	HS159264	yx94f02.s1 Homo sapie	4.52e-175
c	32	215	13.8	445 114	N26159	yx94f02.s1 Homo sapie	4.52e-175
c	33	215	13.8	461 11	H24331	yy54f07.s1 Homo sapie	4.52e-175
c	34	213	13.7	439 63	R61887	yhl0g05.s1 Homo sapie	4.50e-173
c	35	214	13.7	443 5	H05066	y169f06.s1 Homo sapie	4.51e-174
c	36	213	13.7	453 13	H29593	yml6b01.s1 Homo sapie	4.50e-173
c	37	214	13.7	463 130	HS135267	yy20d04.s1 Homo sapie	4.51e-174
c	38	214	13.7	463 116	N53135	yy20d04.s1 Homo sapie	4.51e-174
c	39	212	13.6	349 81	T24036	seq2194 Homo sapiens	4.49e-172
c	40	212	13.6	349 77	T03928	seq2491 Homo sapiens	4.49e-172
c	41	212	13.6	441 72	R92703	yy03c12.s1 Homo sapie	4.49e-172
c	42	212	13.6	495 61	R53536	yyq8f4b07.s1 Homo sapie	4.49e-172
c	43	210	13.5	452 20	H53109	yy08f4c08.s1 Homo sapie	4.46e-170
c	44	210	13.5	499 82	T28553	AB31486F Homo sapiens	4.46e-170
c	45	210	13.5	558 134	HS562300	yy63e01.s1 Homo sapie	4.46e-170

ALIGNMENTS

1 RESULT

ID H587267 standard; RNA; EST; 642 BP.
AC N26587;
DT 30-DEC-1995 (Rel. 46, Created)
DT 30-DEC-1995 (Rel. 46, Last updated, Version 1)
DE yx9le07.s1 Homo sapiens cDNA clone 269124 3' similar to contains
DE Alu repetitive element;
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-642
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevasakis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@watson.wustl.edu High quality sequence stops: 385 Source: IMAGE
CC Consortium, LNL This clone is available royalty-free through LNL
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1140935
FH Key Location/Qualifiers
FT source 1..642
FT /organism="Homo sapiens"
FT /clone="269124"
FT /note="human"
FT mRNA <1..>642
SQ Sequence 642 BP; 182 A; 135 C; 152 G; 166 T; 7 other;

Query Match 33.3%; Score 519; DB 134; Length 642;
Best Local Similarity 97.8%; Pred. No. 0.00e+00;
Matches 544; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

Db 1 gaatctctgatcctttatgattctctgtcacattctgtgagaggatttgggaacc-a 59
|||||
Cp 1522 GATCTCTGATCTCTTATGATTTCTCTGCACATTCTTCTGAGAGGATTTGGGACCCA 1463
|||||
Db 60 agagagcctaattgcataggtacatttaaggcagtcacacaggcagtcacactaaggcag 119
|||||
Cp 1462 AGAGAGCTAAATGGCATAGGTACATTTAAGGCGAGTACACAGGCTACGACTAAGCGAG 1403
|||||
Db 120 tacatagtagctttactatttaaatcctctaccattcttggttagtagaataac 179
|||||
Cp 1402 TACATAGTAGCTTTACTATTATTAATCCTTCTACCAATTTCTGTTAGTAGAATAATAC 1343
|||||
Db 180 ttgtcattagtagatagcattgagcttgaagtgaatgaatgactgogttatat 239
|||||
Cp 1342 TTGCTATTAGTAGATAGCATTTGACCCCTTCAAGTCAATCTAATAGACTGCGCTTATAT 1283
|||||
Db 240 gtcttacctaaagatagatcagccaggcgtggtggctcacgc-tgtattccagcac 298
|||||
Cp 1282 GTCTTACTAGATATAGATATCCAGCGAGCGGTGGCTCGACGCTGTAAATTCAGACAC 1223
|||||
Db 299 ttggaggcgtgagcgccagcatcacctcagggtcaggagtttgagaccagcctgaccaa 358
|||||
Cp 1222 TTGGGAGCGTGGCGGCGGATACCTTGAGTCAGGCTTTGACACAGCGCTGACCAA 1163
|||||
Db 359 catggaataaacccgtctctactaaactacaaatagtcaggcatagtcgacnct 418
|||||

Cp 1162 CATGGAGAACCCCGTCTCTACTAATAACTACAAAATAGTCAGGCATAGTGGCGCAGCC 1103
Db 419 gtttaatcccgactactcaggaggtaagcgaggagattgctgaacctgggagcgag 478
|||||
Cp 1102 TGTATTCCTCCAGCTACTCAGGAGGCTAAGCGAGGAGATGCTTGACCTGGGAGCGGAG 1043
|||||
Db 479 gttatgggtgagcaagatcgccattgcactccagctggcgcaaaaaagcgaactcc 538
|||||
Cp 1042 GTTATGG-TGAGCCAGATCGGCCATTGCTCCAGCTGGCGCAACAAAGCGCAACTC 984
|||||
Db 539 catctcaaaaaaagag 554
|||||
Cp 983 CATCTCAAAAAAAG 968
|||||

RESULT 2
LOCUS N26587 642 bp mRNA EST 29-DEC-1995
DEFINITION yx9le07.s1 Homo sapiens cDNA clone 269124 3' similar to contains
Alu repetitive element;
ACCESSION N26587
NID g1140935
KEYWORDS EST.
SOURCE human clone=269124 primer=ml3-40 forward library=Soares melanocyte
2NBHM vector=pf7T30 (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel=Not I Rsite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGCGAGCGCCAGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pf7T3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 642)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 385
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1140935 Location/Qualifiers
1..642
/organism="Homo sapiens"
/clone="269124"

FEATURES
source

Aug 27 08:30

US-08-231-565A-1.rst

5

/note="human"

<1..>642

BASE COUNT 182 a 135 c 152 g 166 t 7 others

ORIGIN

Query Match 33.3%; Score 519; DB 114; Length 642;

Best Local Similarity 97.8%; Pred. No. 0.00e+00;

Matches 544; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

Db 1 gaatctgtatccttattgatttcttgcacattcttgcagagggatttgggaacc-a 59

Cp 1522 GAATCTGTGATCTTTATGATTTCTTCGCACATTCTTCGAGAGGGATTGGGAACCA 1463

Db 60 agagagctaattgcatagtgatatttaagcagtcacacagggatcagcactaaggcag 119

Cp 1462 AGAGAGCTAATTGCGATAGTACATTAAAGCAGTACACAGGATCAGCAGCTAAGCGAG 1403

Db 120 tacatagtagcttttactatttataatctctaccattcttgcagtttgcagtaaatatc 179

Cp 1402 TACATAGTAGCTTTTACTATTATAATCCTTACCAATTTCTGTTAGTAAATATC 1343

Db 180 ttgtcattagttagaatgacattgagcttgaagtgaatgaaatgaaatgacgttatat 239

Cp 1342 TTGCTATTAGTAGAATAGCATTGACGCTTGAATGAAATGAAATGAAATGAAATGAAAT 1283

Db 240 gtcttaactaatagatagatccagccagggcgtggtggtcaagc-tgtaattccagcac 298

Cp 1282 GTCTTACCTAAGATATAGATCAGCGGCGGTGGCTCAGCCTGTAATTCAGGAC 1223

Db 299 ttggaggctgagggcgagcatcacctgaggtcagggttgcagcagcagcctgaccaa 358

Cp 1222 TTGGGAGGCTGAGCGGGCAGATACCTGAGTGAGGATTGACGACGAGCTGACCAA 1163

Db 359 catggaagaacccctctctactaaaactacaaattagtcaggcatagtcgacacnct 418

Cp 1162 CATGGAAGAACCCGCTCTCTACTAAACTACAAATATAGTCAGGCATAGTGGCGACGC 1103

Db 419 gttatccagctactcaggagctaaaggcaggaattgcttgaacctgggagcggag 478

Cp 1102 TGTATCCAGCTTACTCAGGAGGCTAAGCAGGAGAAATGCTTCACTGGGAGCGGAG 1043

Db 479 gttatgggtgagccaagatcgccattgcaactccagcctgggcaaaaagcgaactcc 538

Cp 1042 GTTATGG-TGAGCGAAGATCGGCCCATTTGCATTCGAGCCTGGGCAACAAAGCAACTC 984

Db 539 catctcaaaaagaag 554

Cp 983 CATCTCAAAAAGAG 968

RESULT 3

LOCUS N23770 593 bp mRNA EST 28-DEC-1995

DEFINITION YX35a12.r1 Homo sapiens cDNA clone 263710 5'.

ACCESSION N23770

NID g1137920

KEYWORDS EST.

SOURCE human clone=263710 primer=T7 library=Soares melanocyte 2NBHM

vector=pT73D (Pharmacia) with a modified polylinker host=DH10B

(ampicillin resistant) Reiter=Not I Reiter2=Eco RI Male. 1st strand

cDNA was primed with a Not I - oligo (dT) primer

[5'-TCTTACCAATCTGAAGTGGGAGCGGCCAGTGTCTTTTCTTTTCTTTTCTTTTCTTTT-3'],

double-stranded cDNA was size selected, ligated to Eco RI adapters

(Pharmacia), digested with Not I and cloned into the Not I and Eco

RI sites of a modified pT73 vector (Pharmacia). Library

Aug 27 08:30

US-08-231-565A-1.rst

6

constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcoteriigii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 593)

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaak, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilton, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

High quality sequence stops: 414

Source: IMAGE Consortium, LIND

This clone is available royalty-free through LIND; contact the

IMAGE Consortium (info@image.lind.gov) for further information.

NCBI gi: 1137920

Location/Qualifiers

source 1..593

/organism="Homo sapiens"

/clone="263710"

/note="human"

<1..>593

BASE COUNT 165 a 140 c 137 g 146 t 5 others

ORIGIN

Query Match 31.3%; Score 488; DB 113; Length 593;
Best Local Similarity 96.5%; Pred. No. 0.00e+00;
Matches 555; Conservative 0; Mismatches 10; Indels 10; Gaps 10;

Db 2 agcagacagaggactctcattaaagagggtcctgtgccctgacctacacagatgccaa 61

Qy 1 AGCAGACAGAGGACTCTCATTAAGGAAGGTCTGCTGCCCTGACCTACAGATGCCAA 60

Db 62 gagaagatgctcactctatctatggttaccaccaagaggggcagcgccactcttacacca 121

Qy 61 GAGAGATGCTCACTTCTATCTATGTTACCCCAAGAGGGCGGCGGCTCTTACACCA 120

Db 122 cgctgaagagggccgtggatcgccatcctgacagtgcctgggagcttactgctca 181

Qy 121 CGGCTGAAGAGGGCCGCTGGATCGGATCTCTGACAGTGCCTGGGAGTCTTACTGCTCA 180

Db 182 tcgctgttggttatgtagaagacgaatggatcacagagccttgatgataaaagtcttc 241

Qy 181 TCGGCTGTGTTGTTAGTAAGACCGAAATGATACAGAGCCTTGTATGATAAAAGCTTC 240

Db 242 atgttggcactcaatgtccttaacaagaagatgcccaagaagggtttgatcatcg 301

Qy 241 ATGTGGCACTCAATGTGCTTTAACAAGAGATGCCCAAGAGGGTTTGTATCATCGGG 300

Db 302 acagcaagtgctcttccaagagaaaaactgtgaacctgtggttcccaatgctccactg 361

301	ACAGCAAAAGTCTCTTTCAAGAGAAAAACTGTGAACCTGTGGTTCCCAATGCTCCACCTG	360
362	cttatgagaactctctgcagaacagtcaccacacaccttatccaccttaagagccagcaga	421
361	CTTATGAGAAACTCTCGCAGACAGTCCACGACACCTTATTTACCTTTAAGAGCCACGCA	420
422	gacacctggagacatctgaaattatttctcctcacactttgttgaaattaatcacgg	481
421	GACACCTG-ACAGACGCTGAAATATTATTCCTC-TCACACTTTTGCTTGAAATTAATACAGA	478
482	ccatctaattggttctccctttggggaatgggtgtaggnaaaatgncacccactctaatna	541
479	C-ATCTAATG-TTCTCC-TTTGG-AATGG-TGTAGGAAAAAATGCAACCATCTCTAATAA	533
542	taaggtcagtggttaaaatttttggtagggccgct	576
534	TAAG-TCAGCT-TTAAAAATTTTAG-TAGTCGCGT-565	

RESULT	4	standard; RNA; EST; 593 BP.
ID	HS770262	
AC	N23770;	
DT	29-DEC-1995 (Rel. 46, Created)	
DT	29-DEC-1995 (Rel. 46, Last updated, Version 1)	
DE	yx35a12.r1 Homo sapiens cDNA clone 263710 5'.	
KW	EST.	
OS	Homo sapiens (human)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.	
RP	[1]	
RN	1-593	
RA	Hillier L., Clark N., Dubouche T., Elliston K., Hawkins M.,	
RA	Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,	
RA	Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,	
RA	Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;	
RA	"The WashU-Merck EST Project";	
RL	Unpublished.	
CC	Contact: Wilson RK WashU-Merck EST Project Washington University	
CC	School of Medicine 4444 Forest Park Parkway, Box 8501,	
CC	MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:	
CC	est@waston.wustl.edu High quality sequence stops: 414 Source: IMAGE	
CC	Consortium, L1NL This clone is available royalty-free through L1NL	
CC	; contact the IMAGE Consortium (info@image.l1nl.gov) for further	
CC	information. NCBI gi: 1137920	
FFH	Key	Location/Qualifiers
FFH	source	1..593
FFT	/organism="Homo sapiens"	
FFT	/clone="263710"	
FFT	/note="human"	
FT	mRNA	<1..>593
FT	Sequence	593 BP; 165 A; 140 C; 137 G; 146 T; 5 other;
SQ	Query Match	31.3%; Score 488; DB 136; Length 593;
	Best Local Similarity	96.5%; Pred. No. 0.00e+00;
	Matches	555; Conservative 0; Mismatches 10; Indels 10; Gaps 10
Ddb	2	agcagacagaggactctcattaaagaaggtgctgcctgccctaccctacaagatgccaa 61
Qy	1	AGCAGACAGAGGACTCTCATTAAGGAAGGTGCTGCTGCCCTGCCCTACCAAGATGCCAA 60
Ddb	62	gagaagatgctcacttcattatcggtaccaccaagaaggggcagcgccactcttacacca 121

	61	GAGAAAGATGCTCACCTTCATCTATGTGTTACCACCAAGAAGGGCAGCGGCCACTCTTACACCA	120
Qy			
Ddb	122	cggctgaagaggccctggatcggcatccctgcactcgacgtgatcctggagtccttaactgctca	181
Qy	121	CGGCTGAAGAGGCCCTGGATCGGCATCCCTGCACGTGATCCTGGGAGTCTTACTGTCTCA	180
Ddb	182	toggctgttgattagttagaagaagaaatggatacacagagccttgatggataaaaagcttctc	241
Qy	181	TGGGCTGTGGTAATTGTAGAAGACGAATAATGGATACAGAGCCTTGATGGATAAAAAGTCTTC	240
Ddb	242	atgtttggcactcaatgtgccttaacaagaagatgccacaagaagggttgatcatcggg	301
Qy	241	ATGTTGGCACTCAATGTGCGCTTAAACAAGAATGCCCCACAGAAAGGGTTGATCATCGGG	300
Ddb	302	acagcaaatgtctctccaagagaaaaactgtgaacctgtggttcccacatgctccacctg	361
Qy	301	ACAGCAAAGTGTCCTTCAAGAGAAAAACTGTGAACCTGTGGTTCCTCAATGCTCCACCTG	360
Ddb	362	cttatgagaacctctctgcagacaagtcaccacacacctattacccttaagagccacga	421
Qy	361	CTTATGAAAACTCTCTGCAGAACAGTCACCAACACCTTTATTCACCTTAAAGAGCCAGCGA	420
Ddb	422	gacacctggagacatgctgaattattctctcctcacacttttgttgaatttaataacagg	481
Qy	421	GACACCTTG-AGACATGCTGAATATTATTTCTC-TCACACTTTTTGCTTGAATTTAATACAGA	478
Ddb	482	ccactctaattggttcccccttgggaatgggtgtagnaaaatgncaaccacctctaatna	541
Qy	479	C-ATCTAATG-TTCTCC--TTTGG-AATGG-TGTAGAGAAAAATGCAAGCCACTCTCTAATAA	533
Ddb	542	taaggtcagtggttaaaatttttgt-agggccgct	576
Qy	534	TAAG-TCAGTG-TTAAAAATTTTAG-TAGGTCGCGCT	565

RESULT	5	standard; RNA; EST; 522 BP.
HS510255		
AC	H97510;	
DT	14-DEC-1995 (Rel. 46, Created)	
DE	14-DEC-1995 (Rel. 46, Last updated, Version 1)	
DE	wllh07.gi Homo sapiens cDNA clone 251965 3' similar to contains	
DE	Alu repetitive element;;	
DE	EST.	
XM		
SW	Homo sapiens (human)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.	
NC	[1]	
RP	1-522	
RA	Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,	
RA	Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,	
RA	Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,	
RA	Travaskis E., Waterston R., Williamson A., Wohldmann P., Willson R.;	
AT	"The WashU-Merck EST Project";	
AT	Unpublished.	
UL		
CC	Contact: Wilson RK WashU-Merck EST Project Washington University	
CC	School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,	
CC	MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:	
CC	estw@stl.wustl.edu High quality sequence stops: 313 Source: IMAGE	
CC	Consortium, LNL This clone is available royalty-free through LNL	
CC	; contact the IMAGE Consortium (info@image.llnl.gov) for further	
CC	information. NCBI gi: 1118395	
CC	Key	Location/Qualifiers
PH		
FT	source	1..522

FT	/organism="Homo sapiens"
FT	/clone="251965"
FT	/note="human"
FT	<1..>522
FT	mRNA
ST	Sequence 522 BP; 143 A; 109 C; 128 G; 135 T; 7 other;
SQ	
	Query Match 30.9%; Score 482; DB 134; Length 522; Best Local Similarity 97.7%; Pred. No. 0.00e+00; Matches 507; Conservative 0; Mismatches 8; Indels 4; Gaps 4;
Db	4 gaatctgatcctttatgatttctctgcacattcttgagaggatttgggaacca 63
Cp	1522 GAATCTGATCCTTTATGATTCTTCGCACATCTTGTGAGGGATTGGGAACCA 1463
Db	64 agagagcataaattgccataggtacatttaaggcagtacacaggcatcacactaaaggcag 123
Cp	1462 ACAGAGCTAAATGGCATAGGTACATTTAAGGCGAGTACACAGGCATCAGACTAAGGCAG 1403
Db	124 tcatatagatctttaocttatattaatactcttcaccaatttctggttagtagaaaaaac 183
Cp	1402 TACATAGTAGCTTTTACTTATTTAAAATCCTTCTACCAATTTCTGTTTTAGTAGAAAAATAC 1343
Db	184 ttgtcattagttagaatgacctgagcccttgaagtgaattgaattagactgcgttat 243
Cp	1342 TTGTCATTACTAGTAATAAGATTGAGCCCTTGAACTGMAATGTAATAGCTGCCGTTAT 1283
Db	244 gcttcactaatgatatagatccagccagcggtggtgcctcacgcctctaattccagcac 303
Cp	1282 GTCTTACCCTAAGATATAGGATCCAGCCAGGCGTGGTGCTCAGCGCTGTATTTCCAGCAC 1223
Db	304 ttttggaggctgagcgcgagatcaectgaagtcaggagttttagaccagcctgacca 363
Cp	1222 TTT- G GGAGGCTGAGGGGGGAGATCCTCAGGTCAGGAGTTTGAGACCAGCCTGACCA 1164
Db	364 acatggngaaaccccgnctctactaaaactacaaaattagtcaggcanagtggcgacn 423
Cp	1163 ACATGG-AGAAGACCCGCTCTACTAANAATAGCAAAATAGTCAGGCATAGTGGCGCAG 1105
Db	424 c-tgtaatccccagctnctcaggaggtcaaggcagaggaattgcttgaacnmgggaggcgg 482
Cp	1104 CTGTAAATCCCGACTACTCAGGAGGCTAAGCGAGAGAAATGCTTGAACCTGGGAGGGG 1045
Db	483 aggttatgmgagcccgatcgcgccattgcaactcag 521
Cp	1044 AGGTTATGGTGAGCCAGCATCGCG-CCATTGCACTCCAG 1007

6	RESULT					
	LOCUS	H97510	522 bp	mRNA	EST	12-DEC-1995
	DEFINITION	yw11h07.s1 Homo sapiens cDNA clone 251965 3' similar to contains				
		Alu repetitive element;.				
	ACCESSION	H97510				
	NID	g1118395				
	KEYWORDS	EST.				
	SOURCE	human clone=251965 primer=m13 -40 forward library=Soares melanocyte 2NDHM vector=pfTT3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Raife1=Not I Rsite2=Eco RI Male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTCAATCTGAAGTCGGAGCGCCGACGTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTT3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.				

Albino.	
ORGANISM	Homo sapiens Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 522)
AUTHORS	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
TITLE	The WashU-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 313 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LINDA; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

	NCBI gi: 1118395					
FEATURES	Location/Qualifiers					
source	1..522					
	/organism="Homo sapiens"					
	/clone="251965"					
	/note="human"					
mRNA	<1..522					
BASE COUNT	143 a	109 c	128 g	135 t		7 others
ORIGIN						

Query Match	30.9%	Score 482;	DB 111;	Length 522;
Best Local Similarity	97.7%	Pred. No. 0.00e+00;		
Matches	507;	Conservative 0;	Mismatches 8;	Indels 4; Gaps 4;
Db	4	gaatctctgatcccttatgattttctctgcacattcttctdgagaggatttgggaaccca	63	
Cp	1572	GAATCTCTGATCCCTTATGATTTCTTCTGCACATCTTCTGTGAGAGGATTTGGGAACCCA	1463	
Db	64	agagagctaaattgccataggtacatttaagcgagctacacagggcatcagcactaaaggcag	123	
Cp	1462	AGAGAGCTAAATTTGCCATAGTAGTACATTTAAGCGAGTACACAGGCATCAGCAGCTTAAGGCAG	1403	
Db	124	tacatagtagcttttaactattttaaatccctctacaaattctcgttttagtagaataatc	183	
Cp	1402	TACATAGTAGCTTTTACTATTATTTAAATCCCTCTACCAATTTCTGCTTTAGTAGAAATATC	1343	
Db	184	ttgcattagttagaatagcattgagccttgaagtgaatgtaattgagactgcgttatat	243	
Cp	1342	TTGCTCAATTAGTTAGATAGCATTTGAGCCTTGAAGTGAATGTAATTAGACTGCGTTATAT	1283	
Db	244	gtcttacctaagatattaggtaccagccagcgctggtgcgtcacgcctgtaattccagcac	303	
Cp	1282	GTCTTACTTAAGATATAGATTCAGCCAGCGCTGGTGGCTCAGCGCTGTAAATTCAGGAC	1223	
Db	304	ttttgggagcctagcgaggcgagatcacctgaaggtcaggaagtttgagaccagcctgacca	363	
Cp	1222	TTT--GGCAGGCTCAGGCGGGGAGATCACCCTGAGCTCAGAGGTTTGACACAGCGTCACCA	1164	

Db 364 acatggnngaaccnctactactaaactacaaataatagtcaggcanagtgccgacn 423
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1163 ACATGG-AGAACCCCGTCTCTACTTAAACTACAAAATAGTCAGGCATAGTGGCGCAG 1105
Db 424 c-tgtaatccagctnctcaggaggtaagcaggagaattgcttgaacnnggagcg 482
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1104 CCTGAATCCAGCTACTACGAGGCTAAGCGAGGAGAAATTCCTTGAACCTGGGAGCGG 1045
Db 483 aggttatgngagccagatcgcggccatgcactccag 521
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1044 AGGTTATGCTGCGCAAGATCGCG-CCATTGCCTCCAG 1007

RESULT 7

LOCUS H97498 540 bp mRNA EST 12-DEC-1995
DEFINITION yw11f07.s1 Homo sapiens cDNA clone 251941 3' similar to contains
Alu repetitive element;

ACCESSION H97498

NID g1118383

KEYWORDS EST.

SOURCE human clone=251941 primer=ml3 -40 forward library=Soares melanocyte
2NHM vector=pT73D (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo (dT) primer
[5'-TCTTACCAATCTGAGTGGGCGCGCGCGAGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonalido. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygia; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 540)

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 377

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

FEATURES

source

NCBI gi: 1118383

Location/Qualifiers

1..540

/organism="Homo sapiens"

/clone="251941"

/note="human"

mRNA <1..>540
BASE COUNT 153 a 111 c 135 g 136 t 5 others
ORIGIN

Query Match 30.0%; Score 468; DB 111; Length 540;
Best Local Similarity 96.8%; Pred. No. 0.00e+00;
Matches 520; Conservative 0; Mismatches 9; Indels 8; Gaps 7;

Db 3 gaatctgatcctttatgtattctctgcacattctgtgagaggatttgggaacca 62
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1522 GAATCTCTGATCCTTTATGATTTCTTCGCACATTTCTGTGAGAGGATTTGGGACCCA 1463
Db 63 agagagctaaattgccatagtcatttaaggcagtcacacaggcatcagcactaaggcag 122
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1462 AGAGAGCTAAATTTGCCATAGGTACATTTTAAAGGCAGTACACAGGCATCAGCCTAAGGCAG 1403
Db 123 tacatagtagctttactatttaatactctaccattctcggtttagtagaataac 182
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1402 TACATAGTAGCTTTTACTTTATTTAAATCCTTCTACCAATTTCTGGTTAGTAGAAAATAC 1343
Db 183 ttgtcattagtagaataagcattgagcctgaagtgaatgaattagactgcgttatan 242
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1342 TTGCTATTAGTATAGATAGCATTTGAGCCCTTGAGTGCATTAATGTAATTAGACTGCGTTATAT 1283
Db 243 gtcttacctaatagatagatccagccaggcgtggtgcctcacgcctgaattccagcac 302
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1282 GTCTTACCTAAGATATAGGATCCAGCCAGGCGGTGGCTCAGCCCTGTAAATTCAGGCAC 1223
Db 303 ttggaggagctgagcggcgagatcacctgaggtcagantttgagaccagcctggacc 362
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1222 TTTGGAGGCTGAGSGCGGAGATCACCTG-AGGTCAGGAGTTTGAGACCAGCCTG-ACC 1165
Db 363 aacatggagaaccctgctctactaaactacaaaattagtcaggcatagtgccgacag 422
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1164 AACATGGAGAAACCCCGCTCTACTTAAAACTACAAAATTAGTCAGGCATAGTGGCGCAGG 1105
Db 423 c-tgtatccacagctactcaggaggctgaagcaggagaattgcttgaacctggggaagcg 481
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1104 CCTGTAATCCAGCTACTCAGGAGGCTAAGCGAGGAGTAATGCTTGACCTGCGG--AGCG 1047
Db 482 cgaaggttaagggtgagccagatcgccantgcacttcagctcagctggggnaaaaaa 538
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Cp 1046 -CGAGGTTATGG-TGAGCCAAAGATCGCCCATTTGCACCT-CCAGCCTGGGCAACAAA 993

RESULT 8

ID H5498253 standard; RNA; EST; 540 BP.
AC H97498;
DT 14-DEC-1995 (Rel. 46, Created)
DT 14-DEC-1995 (Rel. 46, Last updated, Version 1)
DE yw11f07.s1 Homo sapiens cDNA clone 251941 3' similar to contains
DE Alu repetitive element;
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN [1]
RP 1-540
RA Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
RA Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
RA Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
RA Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., Wilson, R.;
RT "The WashU-Merck EST Project";
RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@watson.wustl.edu High quality sequence stops: 377 Source: IMAGE
CC Consortium, L1NL This clone is available royalty-free through L1NL
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1118383
FH Key Location/Qualifiers
FH Key
FT source 1..540
FT /organism="Homo sapiens"
FT /clone="251941"
FT /note="human"
FT mRNA
FT <1..>540
SQ Sequence 540 BP; 153 A; 111 C; 135 G; 136 T; 5 other;

Query Match 30.0%; Score 468; DB 133; Length 540;
Best Local Similarity 96.8%; Pred. No. 0.00e+00;
Matches 520; Conservative 0; Mismatches 9; Indels 8; Gaps 7;

Db 3 gaatctgacatttattgattctctgacattctctgagggatttgggaacca 62
|||||
Cp 1522 GAATCTGTATGCTTTATGATTTCTTCTGCACATCTTTGTGAGAGGGATTGGCAACCA 1463

Db 63 agagagtaattgccatagttacatttaaggcaggtacacaggcatcagcactaaggcag 122
|||||
Cp 1462 AGAGAGCTAAATTTGGCATAGGTACATTTTAAGGCGGTACACAGGCATCAGCAATAGGCAG 1403

Db 123 tacatagtagctttacttatttaaatcctctaccattctcgtttagtagtaaaatc 182
|||||
Cp 1402 TACATAGTAGCTTTTACTTATTAAATCCTTCTACCAATTTCTGGTTAGTAGAAATATC 1343

Db 183 ttgtcattagtagtagcattgagccttgaagtgaatgaatgactgcgttatan 242
|||||
Cp 1342 TTGCTATTAGTAGTAGATTTGAGCCTTGAAGTGAATGTAATTTAGACTGCGTTATAT 1283

Db 243 gtcttacctaaagatataagatccagcaggcgtggtggctcacgcctgtaattccagcac 302
|||||
Cp 1282 GTCTTACTTAAGATATAGGATACAGCCAGCGGTGGCTCAGCCTGTAATTTCCAGCAC 1223

Db 303 ttggaggctgagcgggagacataccctggaggtcaggantttgagaccagcctggacc 362
|||||
Cp 1222 TTTGGAGGCTGAGGCGGCGACATCACCTG-AGGTCAGGAGTTTGAGACCGCTG-ACC 1165

Db 363 aacatggagaaacccgtctctactaaaactacaaaattagtcaggcatagtgccgcacy 422
|||||
Cp 1164 ACATGAGAAACCCGCTCTCTACTTAAACTTCAAAAATTAGTCAGGCATAGTGGCGCAG 1105

Db 423 c-tgtlaatccagctactcaggaggctaagcgaggaggaattgcttgaacctggggaagcc 481
|||||
Cp 1104 CTTGTAATCCAGCTACTCAGGAGGCTAAGCGAGGAGAAATTGCTGAACCTGGG--AGGC 1047

Db 482 cggaggttaagggtgagccagatcgcccaantgcaactccagntcggggnaaaaaa 538
|||||
Cp 1046 -CGAGTTATGG-TGAGCAAGATCGGCCATTGCACT-CCAGCCTGGGCAACAAA 993

RESULT 9
LOCUS N20091 473 bp mRNA EST 15-DEC-1995
DEFINITION yx36a12.s1 Homo sapiens cDNA clone 263806 3'.
ACCESSION N20091
NID g1124758
KEYWORDS EST.
SOURCE human clone=263806 primer=m13 --40 forward library=Soares melanocyte

2NBHM vector-pT7T3D (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel-Not I Reite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo (dT) primer
[5'-TGTTACCAATCTGAGTGGAGCGCGCGAGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM
Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 473)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 363
Source: IMAGE Consortium, L1NL
This clone is available royalty-free through L1NL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1124758 Location/Qualifiers
1..473
/organism="Homo sapiens"
/clone="263806"
/note="human"
<1..>473

BASE COUNT 127 a 97 c 97 g 151 t 1 others

ORIGIN

Query Match 28.3%; Score 441; DB 112; Length 473;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 467; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 catgattgcaacatttatgaggagtttcccaatttaataatttcctcctc 60
|||||
Cp 642 CATGATTCGCACATTTATTCATGGAGTTTCCCAATTTAATATTCATCATCTTCCTC 583

Db 61 acatgattagtagctgctagcgagcctactaaaatttaacactgacttatattagagat 120
|||||
Cp 582 ACATGATTAGTACTGCTAGGAGACCTACTATAAATTTTAACACTGACTTATTATAGAGAT 523

Db 121 ggcctgcatctttctcacattccaaaggagacattagatgctctgtattaaattcaa 180
|||||
Cp 522 GCGTTCGATTTTCTTACACCATTCACCAAGGAGACATTAGATGCTGTATTAAATTCAA 463

Db 181 gcaaaagtgtgagagaataatttcagcatgctcaggtgtcctcgctggctcttaaggtg 240
|||||

Cp 462 GCAAAAGTGTGAGAGAAATAATTTCAGCATGTCTCAGGTGTCTCGCTGGCTTAAAGGTG 403
Db 241 gataagtggtggtgactgtttctgagagagtttctcataagcaggtgagcattgggaa 300
Cp 402 AATAAGGTGGTGTGCTGACTTCTTCAGAGAGTCTTCTCATAGCAGGTGGAGCATTTGGAA 343
Db 301 ccacaggtccacagttttctcttgagagacactttgctgccgatgatcaaacctt 360
Cp 342 CCACAGGTTACAGTTTTCTCTTGAAGAGACACTTGTCTGCCGATGATCAAAACCTT 283
Db 361 ctgtgggcatcttctgttaaggccacattngagtgccaacacagcgggtgagcag 420
Cp 282 CTGTGGGCACTCTCTTTAAGG-CACATT-GAGTGCAC-ATG-AGACTTTTATCC 227
Db 421 atcaaggtctgtatccattctcttacaataccaacacagcgggtgagcag 473
Cp 226 ATCAAGGCTGTATCCATTCTCTTCAATAACCAACAGCGATGAGCAG 174

RESULT 10

ID H5091255 standard; RNA; EST; 473 BP.
AC N20091;
DT 17-DEC-1995 (Rel. 46, Created)
DT 17-DEC-1995 (Rel. 46, Last updated, Version 1)
DE yx36a12.s1 Homo sapiens cDNA clone 263806 3'.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-473
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Holtman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevasakis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@washington.wustl.edu High quality sequence stops: 363 Source: IMAGE
CC Consortium, IMAGE This clone is available royalty-free through IMAGE
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1124758
FH Key Location/Qualifiers
FT source 1..473
FT /organism="Homo sapiens"
FT /clone="263806"
FT /note="human"
FT mRNA <1..>473
SQ Sequence 473 BP; 127 A; 97 C; 97 G; 151 T; 1 other;

Query Match 28.3%; Score 441; DB 129; Length 473;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 467; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 catgcatgcaacatttatgtgagtttcccaatttaatttctcatcttcctc 60
Cp 642 CATGATGCAACATTTATTCATGAGTTTCCCAATTTATTTCTCATCATTTCTCCTC 583
Db 61 acatgattagctgctagcgagcactactaaattttaacactgacttattattagagat 120
Cp 582 ACATGATTAGTACTGCTAGCGGACCTACTAAATTTTAACTGACTTATTATTAGAGAT 523

Db 121 ggcttcatttttctacacattccaagaagagacattatgatgtctgtattaaattcaa 180
Cp 522 GGTTCGATTTTCTCAGCATTCACCATTCAGAGAGACATTAGATGTCTGTATTAATTCAA 463
Db 181 gcaaaagtgtgagagaaataatttcagcatgtctcaggtgtctcgtcgctcttaaggtg 240
Cp 462 GCAAAAGTGTGAGAGAAATAATTTCAGCATGTCTCAGGTGTCTCGCTGGCTTAAAGGTG 403
Db 241 gataagtggtggtgactgtttctgagagagtttctcataagcaggtgagcattgggaa 300
Cp 402 AATAAGGTGGTGTGCTGACTTCTTCAGAGAGTCTTCTCATAGCAGGTGGAGCATTTGGAA 343
Db 301 ccacaggtccacagttttctcttgagagacactttgctgccgatgatcaaacctt 360
Cp 342 CCACAGGTTACAGTTTTCTCTTGAAGAGACACTTGTCTGCCGATGATCAAAACCTT 283
Db 361 ctgtgggcatcttctgttaaggccacattngagtgccaacacagcgggtgagcag 420
Cp 282 CTGTGGGCACTCTCTTTAAGG-CACATT-GAGTGCAC-ATG-AGACTTTTATCC 227
Db 421 atcaaggtctgtatccattctcttacaataccaacacagcgggtgagcag 473
Cp 226 ATCAAGGCTGTATCCATTCTCTTCAATAACCAACAGCGATGAGCAG 174

RESULT 11

LOCUS N24677 433 bp mRNA EST 28-DEC-1995
DEFINITION yx91a09.s1 Homo sapiens cDNA clone 269080 3'.
ACCESSION N24677
NID g1138827
KEYWORDS EST.
SOURCE human clone=269080 primer=ml3-40 forward library=Soares melanocyte
2Nblm vector=pT7T3D (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo (dT) primer
{5'-TGTTACCAATCTGAAGTGGGCGCGCCGACAGTTTTTTTTTTTTTTT-3'},
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 433)

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Holtman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

High quality sequence stops: 315
Source: IMAGE Consortium, L1NL
This clone is available royalty-free through L1NL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1138827 Location/Qualifiers
source 1..433
/organism="Homo sapiens"
/clone="269080"
/note="human"
mRNA <1..>433
BASE COUNT 119 a 85 c 86 g 142 t 1 others
ORIGIN

Query Match 27.1%; Score 423; DB 113; Length 433;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 430; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 atcatnctgcaacattattgatggagtttcccaatttaatttctcatcattcc 60
|||||
Cp 644 ATCATGCAATGCAACATTATTGATGGAGTTTCCCAATTAATTTCTCATTTCC 585
|||||
Db 61 tcacatgattagtagtctagcgacctactaaaattttaaacactgactattattagag 120
|||||
Cp 584 TCACATGATTAGTAGTCTAGCGGACCTACTAAAAATTTTAACACTGACTTATTATTAGAG 525
|||||
Db 121 atggttgcatttttctcaaccattccaaaggagaaacattagatgtctgtattaaattc 180
|||||
Cp 524 ATGGCTGCATTTTCTTACACCATTTCCAAAGGAGAACATTAGATGCTGTATTAAATTC 465
|||||
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgctggtcttaagg 240
|||||
Cp 464 AAGCAAAAGTGTGAGAGAAATTAATTCAGCATGCTCTCAGGTGTCTCGCTGCTTTAAGG 405
|||||
Db 241 tgaataaagggtggtgactgttctcgagagagtttctcataaagcaggtggagcattggg 300
|||||
Cp 404 TGAATAAGGTGGTGGTCACTGTTCTCGACAGAGTTTCTCATTAAGCAGGTGGAGCATTTGGG 345
|||||
Db 301 aaccacaggttcaacagtttttctcttgaagagacactttgctcccgatgatcaaac 360
|||||
Cp 344 AACCAAGGTTCAACAGTTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-CATCAAAAC 286
|||||
Db 361 cctcttggggcatcttctgttaaggcacattgagtgcacacatgaagactttttatcca 420
|||||
Cp 285 CTTCCTTGGGGCATCTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-CATCAAAAC 226
|||||
Db 421 tcaaggtctctgta 433
|||||
Cp 225 TCAGGCTCTGTGA 213
|||||

RESULT 12
ID HS677260 standard; RNA; EST; 433 BP.
AC N24677;
DT 30-DEC-1995 (Rel. 4.6, Created)
DT 30-DEC-1995 (Rel. 4.6, Last updated, Version 1)
DE yx91a09.s1 Homo sapiens cDNA clone 269080 3'.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-433

RA Hillier L., Clark N., Duboue T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaekis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@wustl.wustl.edu High quality sequence stops: 315 Source: IMAGE
CC Consortium, L1NL This clone is available royalty-free through L1NL
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1138827

FH Key Location/Qualifiers

FH source 1..433
/organism="Homo sapiens"
/clone="269080"
/note="human"
FT mRNA <1..>433
SQ Sequence 433 BP; 119 A; 85 C; 86 G; 142 T; 1 other;

Query Match 27.1%; Score 423; DB 135; Length 433;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 430; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 atcatnctgcaacattattgatggagtttcccaatttaatttctcatcattcc 60
|||||
Cp 644 ATCATGCAATGCAACATTATTGATGGAGTTTCCCAATTAATTTCTCATTTCC 585
|||||
Db 61 tcacatgattagtagtctagcgacctactaaaattttaaacactgactattattagag 120
|||||
Cp 584 TCACATGATTAGTAGTCTAGCGGACCTACTAAAAATTTTAACACTGACTTATTATTAGAG 525
|||||
Db 121 atggttgcatttttctcaaccattccaaaggagaaacattagatgtctgtattaaattc 180
|||||
Cp 524 ATGGCTGCATTTTCTTACACCATTTCCAAAGGAGAACATTAGATGCTGTATTAAATTC 465
|||||
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgctggtcttaagg 240
|||||
Cp 464 AAGCAAAAGTGTGAGAGAAATTAATTCAGCATGCTCTCAGGTGTCTCGCTGCTTTAAGG 405
|||||
Db 241 tgaataaagggtggtgactgttctcgagagagtttctcataaagcaggtggagcattggg 300
|||||
Cp 404 TGAATAAGGTGGTGGTCACTGTTCTCGACAGAGTTTCTCATTAAGCAGGTGGAGCATTTGGG 345
|||||
Db 301 aaccacaggttcaacagtttttctcttgaagagacactttgctcccgatgatcaaac 360
|||||
Cp 344 AACCAAGGTTCAACAGTTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-CATCAAAAC 286
|||||
Db 361 cctcttggggcatcttctgttaaggcacattgagtgcacacatgaagactttttatcca 420
|||||
Cp 285 CTTCCTTGGGGCATCTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-CATCAAAAC 226
|||||
Db 421 tcaaggtctctgta 433
|||||
Cp 225 TCAGGCTCTGTGA 213
|||||

RESULT 13
ID HS796259 standard; RNA; EST; 396 BP.
AC N24796;
DT 30-DEC-1995 (Rel. 4.6, Created)
DT 30-DEC-1995 (Rel. 4.6, Last updated, Version 1)

DE yx97f10.s1 Homo sapiens cDNA clone 269707 3'.

KW EST.

OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.

RN [1]

RP 1-396
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevasik E., Waterston R., Williamson A., Wohlmann P., Wilson R.,
RT "The WashU-Merck EST Project";
RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@wustl.edu High quality sequence stops: 310 Source: IMAGE
CC Consortium, L1NL This clone is available royalty-free through L1NL
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1138946

FH Key Location/Qualifiers

FT source 1..396 /organism="Homo sapiens"

FT /clone="269707"

FT /note="human"

FT mRNA <1..>396

FT Sequence 396 bp; 108 A; 75 C; 80 G; 131 T; 2 other;

Query Match 25.3%; Score 394; DB 136; Length 396;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atcatnctgcaacatttattgatggagtttcccaatttaatttctcatcttcc 60

|||||

Cp 644 ATCATGCAATTCGACATTAATGATGGAGTTTCCCAATTTAATATTCATCATTTCC 585

|||||

Db 61 tcacatgattagctgctagcggaacctactaaattttaacactgacttatttagag 120

|||||

Cp 584 TCACATGATTAGTACTGCTAGCGGACCTACTAAATTTTAACTGACTTATTATTAGAG 525

|||||

Db 121 atggcttgcatcttctacacattcccaaggagacattagatgtctgtattaatcc 180

|||||

Cp 524 ATGGCTTCGATTTTCTTACACATTCCTCAAGAGGAGACATTAGATGCTGTATTAATTC 465

|||||

Db 181 aagcaaaagtgtgagagaataatttcagcagtgctcaggtgtctcgctggtcttaagg 240

|||||

Cp 464 AGCAAAAGTGTGAGAGAATAATTTACGATGCTTCAGGTGCTCCGCGCTCTTAAGG 405

|||||

Db 241 tgaataaagtggtggtgactgttctgcagagtttctcataaagcaggtgagcattggg 300

|||||

Cp 404 TGAATAAGGTGGTGGTGAAGTCTGTCGACAGAGTCTCTATAAGAGGGTGCAGCATTTGGG 345

|||||

Db 301 aacccagggttcacagttttcttcttgagagacatttctgtcccgatgataaaccc 360

|||||

Cp 344 AACCAAGGTTACAGATTTTCTTCTTGAAGAGACACTTTGCTGTCGCGGATGATCAAAACC 285

|||||

Db 361 ttcttggtggcactcttcttgaaggacattgag 395

|||||

Cp 284 TTCTTGTGGGATCTTCTTGTGTAAGGCACATTGAG 250

|||||

RESULT 14

LOCUS N24796 396 bp mRNA EST 28-DEC-1995

DEFINITION yx97f10.s1 Homo sapiens cDNA clone 269707 3'.

ACCESSION N24796

NID g1138946

KEYWORDS EST.

SOURCE human clone=269707 primer=ml3 -40 forward library=Soares melanocyte
2N8M vector=pt7T3D (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTACCAATCTCAAGTGGAGCGCGCGAGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pt7T3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 396)

REFERENCE

AUTHORS

Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
Trevasik E., Waterston R., Williamson A., Wohlmann P. and
Wilson R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 310
Source: IMAGE Consortium, L1NL

This clone is available royalty-free through L1NL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1138946

Location/Qualifiers

1..396

/organism="Homo sapiens"

/clone="269707"

/note="human"

<1..>396

BASE COUNT 108 a 75 c 80 g 131 t 2 others

ORIGIN

Query Match 25.3%; Score 394; DB 113; Length 396;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atcatnctgcaacatttattgatggagtttcccaatttaatttctcatcttcc 60

|||||

Cp 644 ATCATGCAATTCGACATTAATGATGGAGTTTCCCAATTTAATATTCATCATTTCC 585

|||||

Db 61 tcacatgattagctgctagcggaacctactaaattttaacactgacttatttagag 120

|||||

Cp 584 TCACATGATTAGTACTGCTAGCGGACCTACTAAATTTTAACTGACTTATTATTAGAG 525

|||||

Db 121 atggcttgcatcttctacacattcccaaggagacattagatgtctgtattaatcc 180

Cp 524 ATGGCTTGCATTTTCTACACCATCCAAAGGAGAACATTAGATGCTGTATTAATTC 465
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgtggtcttaagg 240
Cp 464 AAGCAAAAGTGTGAGAGAAATAATTTCAGCATGCTCTCAGGTGTCTCGCTGGCTCTTAAGG 405
Db 241 tgaataaagtggtggtgctgttcgagagagatttctcataagcagggtgagcattggg 300
Cp 404 TGAATAAGGTGGTGTGCTGCTGCAGACAGTTCCTATAGCAGGTGGAGCATTTGGG 345
Db 301 aaccacaggttcacagttttcttcttgagagacactttgctgtcccgatgatcaaaccc 360
Cp 344 AACCCAGGTTACAGTTCCTTCTTGAAGAGACACTTTGCTGCCGATGATCAAAACC 285
Db 361 ttctgtgggcattcttcttgaaggcacattgag 395
Cp 284 TTCTGTGGCATCTCTTGTGAAGCACAATTGAG 250

RESULT 15

LOCUS N26562 401 bp mRNA EST 29-DEC-1995
DEFINITION Yx62h05.s1 Homo sapiens cDNA clone 266361 3'.
ACCESSION N26562
NID q1140910
KEYWORDS EST.
SOURCE human clone=266361 primer=ml3 -40 forward library=Soares melanocyte
2NBHM vector=pT7T30 (pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel=Not I Rsite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTACCAATCTGAAGTGGAGCGCGCGAGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 401)

REFERENCE

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfs, T., Soares, M., Tan, F.,
Trevas, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

COMMENT

Unpublished (1995)
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 305
Source: IMAGE Consortium, LIND
This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.lind.gov) for further information.

NCBI gi: 1140910

FEATURES
source Location/Qualifiers
1..401
/organism="Homo sapiens"
/clone="266361"
/note="human"
<1..>401
BASE COUNT 110 a 79 c 82 g 129 t 1 others
ORIGIN
Query Match 25.1%; Score 392; DB 114; Length 401;
Best Local Similarity 99.5%; Pred. No. 0.00e+00;
Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 1 atcgtcattgcacatttattgatggaggtttcccaatttataatttctcatcttcc 60
Cp 644 ATCATGCATTGCAACATTTATTGATGGAGTTTCCCAATTTTAATATTTTCATCATTTCC 585
Db 61 tcacatgattagtagctagcgacctactaaattttaacacactgaattattattagag 120
Cp 584 TCACATGATTAGTACTGCTACGCGACTACTAAATTTTAACACTGACTTATTATTAGAG 525
Db 121 atggtcttgcatttttctacacattccaaaggagacattagatgtctgtattaaattc 180
Cp 524 ATGGCTTGCATTTTCTACACCATTTCCAAAGGAGAACATTAGATGCTGTATTAAATTC 465
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgtggcnccttaagg 240
Cp 464 AAGCAAAAGTGTGAGAGAAATAATTTCAGCATGCTCTCAGGTGTCTCGCTGGCTCTTAAGG 405
Db 241 tgaataaagtggtggtgctgttcgagagagtttctcataagcagggtgagcattggg 300
Cp 404 TGAATAAGGTGGTGTGCTGCTGCAGACAGTTCCTCATAGCAGGTGGAGCATTTGGG 345
Db 301 aaccacaggttcacagtttttctcttgagagacactttgctgtcccgatgatcaaaccc 360
Cp 344 AACCCAGGTTACAGTTCCTTCTTGAAGAGACACTTTTGTGTCGCCGATGATCAAAACC 285
Db 361 ttctgtgggcattcttcttgaaggcacattgag 395
Cp 284 TTCTGTGGCATCTCTTGTGAAGCACAATTGAG 250

Search completed: Tue Aug 27 08:43:15 1996

Job time : 642 secs.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	RESULT	1			
	LOCUS	HU06452	1524 bp	mRNA	PRI 25-JUN-1994
	DEFINITION	Human melanoma antigen recognized by T-cells (MART-1) mRNA.			
	ACCESSION	U06452			
	NID	q476131			

REFERENCE 1 (bases 1 to 1524)
AUTHORS Kawakami, Y., Elyahu, S., Delgado, C.H., Robbins, P.F., Rivoltini, L., Topalian, S.L., Miki, T. and Rosenberg, S.A.
TITLE Cloning of the gene coding for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91, 3515-3519 (1994)
MEDLINE 94224770
REFERENCE 2 (bases 1 to 1524)
AUTHORS Kawakami, Y.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-1994) Yutaka Kawakami, Surgery Branch, National Cancer Institute, National Institutes of Health, 9000 Rockville Pike, Bldg. 10, Rm. 2B42, Bethesda, MD 20892, USA
COMMENT NCB1 gi: 476131
FEATURES
source Location/Qualifiers
1..1524
/organism="Homo sapiens"
/cell_line="melanoma cell line, 501mel"
/tissue_type="melanoma"
54..410
/gene="MART-1"
/standard_name="melanoma antigen recognized by T-cells"
/note="NCBI gi: 476132"
/codon_start=1
/db_xref="PID:9476132"
/translation="MPREDAHFYGYPKKGHSYVTAEEAAGIGILTVILGVLLIG
CWYCHRRNGYBALMDKSLHVGTCALTRPCQEGFDHRDQSVLQKNCPEVVPNAPP
ATKELISAEQSPPPYSP"
BASE COUNT 435 a 330 c 324 g 435 t
ORIGIN
Query Match 97.8%; Score 1524; DB 52; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 acgacagaggactctcattaggaggtgtcctgtgccctgacctacagatgccaa 60
QY 1 AGCAGACAGGACTCTCATTAAGGAAGGTGCTGCTGCGCTGCCTCAGAGATGCCAA 60
Db 61 gagaagatgctcactctatcattgtaaccccaagaaggcgccagccactctacacca 120
QY 61 GAGAAAGTCTCATTCTATGTTATGCCCAAGAGGGGCGAGCCACTCTTACACCA 120
Db 121 cggctgaagaggccgtgggtcggatcctgacagtgatcctggagtgcttactgetca 180
QY 121 CGGCTGAAGAGGCCGCTGGGATCGGCATCTTGACATGTCCTGGGAGTCTTACTGCTCA 180
Db 181 tcggctgttggtattgtgaagacgaaatggatagagcccttgatggataaaagtcttc 240
QY 181 TCGGCTGTGTTGATTGTGAAGACGAAATGGATACAGAGCCCTTGATGATAAAGTCTTC 240
Db 241 atgttggcactcaatgtgcttaacgaagatgcccacaagaagggtttgatcatcggg 300
QY 241 ATGTTGGCACTCAAATGTCCTTAAAGAAGATGCCCAAGAGGGGTTTGATCATCGGG 300
Db 301 acagcaagtgctctctcaagagaacactggaacctgtggttcccaatgctccacctg 360
QY 301 ACAGCAAGTGTCTCTTCAAGAGAAAACGTGAACCTGTGCTTCCCAATGCTCCACCTG 360
Db 361 cttatgagaactctctgagaacagctcaccacaccttattcactctaaagagcagcga 420
QY 361 CTTATGAGAACTCTCTGACAGACAGTCAACACCCACTTATTCACCTTAAGAGCCGCA 420

Db 421 gacacctgagacagtgcgaattatttctctcacacttttctgcttgaatttaataacagaca 480
QY 421 GACACCTGAGACAGTGCGAATTATTTCTCTCACACTTTTCTGCTGAATTTAATAACAGACA 480
Db 481 tctaattgtctctcttggaaatggtgtaggaaaaatgcaagccatctcttaataaagtca 540
QY 481 TCTAATGTTCTCTCTTGGAAATGCTGAGGAAAATGCAAGCCATCTCTAATAAAGTCA 540
Db 541 gtgttaaaattttagtagtcgctgacgtactaatcatgtgaggaatgatgagaat 600
QY 541 GTGTTAAAAATTTAGTAGGTCGCTAGCAGTACTTAATCATGTGAGGAAATGATCAGAAAT 600
Db 601 attaaatgggaaaaactccatcaataatgttgcaatgcacatgatactatctgtgccagag 660
QY 601 ATTAAATTGGGAAAATCCATCAATAAATGTTGCAATGCATGATCTATCTGTGCCAGAG 660
Db 661 gtaatgttagtaaatccatggtgttattttctgagagacagaattcaagtggttattctg 720
QY 661 GTAATGTTAGTAATCCATGCTGTATTTTCTGAGAGACAGATTCACAGTGGCTATTCTG 720
Db 721 gggccatccaatctcttacttgaattggcctaataacaaactagtcagggttttcca 780
QY 721 GGCCCATCAAAATTTCTTTACTTGAATTTGGCTAATAACAACACTAGTCAGGTTTTCGA 780
Db 781 acctgaccgacatgaactgtcacagaattgttccagtactatggagtgctcacaaag 840
QY 781 ACCTTGACCGACATGAACGTACACAGAAATGTTTCCAGTACTATGGAGTGTCTACAAGG 840
Db 841 atactttacaggttaagaacaaagggtgactggcctatttactatcgtacaagaacatgctc 900
QY 841 ATACTTTTACAGGTTAAGACAAAGGGTTGACTGGCCTATTTATCTGATCAAGAACATGTC 900
Db 901 agcaatgtctcttgtgctctaaattcttattactacaataatatttgaagatccc 960
QY 901 AGCAATGCTCTTTGCTGCTTAAATTTCTATTATACATAATATATTGTAAGATCC 960
Db 961 tatagctcttttttttgagatggaggttgcctttgttggccaggctggagtgcaatgg 1020
QY 961 TATAGCTCTTTTTTTTGAGATGGAGTTTCGCTTTTGTGCCAGGCTGGAGTGCAATGG 1020
Db 1021 cggcatcttggctcaaccataaacctccgctcccagggtccaagcaattctcctgacctagc 1080
QY 1021 CGGATCTTGGCTCACATAACCTCCGCTCCAGGTTCAAGCAATTTCTCTGCTTAGC 1080
Db 1081 ctectgagtagctgggattacaggcgtgcgcacactatgctgacactaattttttagtttta 1140
QY 1081 CTCTCTGAGTACTGGGATTACAGGCTGGCGCACTATGCTGACTAAATTTTGTAGTTTAA 1140
Db 1141 gtgagacggggtttctccatgttggctcaggctggctcctcaactcctgacctcagggtat 1200
QY 1141 GTGAGACGGGGTTTCTCCATGTTGGTCAGGGCTGGTCTCAAACTCTCTGACCTCAGGTGAT 1200
Db 1201 ctgccgctcagcctcccaaatgctgggaattacaggcgtgagccaccacgcctggctg 1260
QY 1201 CTGCCGCTCAGCCTCCCAAAAGTGTGGAAATTACAGGCGGTGAGCCACCGCCTGGCTG 1260
Db 1261 gatcctatatcttaggtgaagacataaacgcaggtcttaattacatttcacttcaaggctca 1320
QY 1261 GATCCTATATCTTAGGTGAACACATAAACGCAGTCTTAATTACATTTCACTTCAAGGCTCA 1320
Db 1321 atgctattcttaactaatgacaagtattttctactaaaccagaataatggtagaagattta 1380
QY 1321 ATGCTATTCTTAACATAATGACAAGTAATTTTCTACTAAACCCAGAAATTTGGTAGAGGATTA 1380

```

exon      238..334
           /number=3
exon      335..448
           /number=4
exon      449..752
           /number=5
polyA signal 730..735
polyA_site 752
           /note="8 A residues"
BASE COUNT 239 a 170 c 163 g 180 t

```

Eukaryotae: Hyperchondria; Eukaryote crown group; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

Query Match 15.3%; Score 239; DB 68; Length 291;
Best Local Similarity 92.1%; Pred. No. 2.14e-104;
Matches 267; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

RX MEDLINE; 8219044.
 RA Groffen J., Heisterkamp N., Grosveld F., de Ven W.,
 RA Stephenson J.R.;
 RT "Isolation of human oncogene sequences (v-fes homolog) from a
 RT cosmid library";
 RN Science 216:1136-1138(1982).
 RX [3]
 RX MEDLINE; 91329820.
 RA Morris C.M., Heisterkamp N., Groffen J., Fitzgerald P.H.;
 RT "Entire ABL gene is joined with 5'-BCR in some patients with
 RT Philadelphia-positive leukemia";
 RN Blood 78:1078-1084(1991).
 RX [4]
 RX 27793-28338, 28792-29265
 RX MEDLINE; 87028219.
 RA Shitvelman E., Lifshitz B., Gale R.P., Roe B.A., Canaani E.;
 RT "Alternative splicing of RNAs transcribed from the human abl gene
 RT and from the bcr-abl fused gene";
 RN Cell 47:277-284(1986).
 RX [5]
 RX 27108-29899
 RX MEDLINE; 88038877.

RA Bernards A., Rubin C.M., Westbrook C.A., Paskind M., Baltimore D.;
RT "The first intron in the human c-abl gene is at least 200
RT kilobases long and is a target for translocations in chronic
RT myelogenous leukemia";
RL Mol. Cell. Biol. 7:3231-3236(1987).
RN [6]
RX MEDLINE; 82150208.
RA Mount S.M.;
RT "A catalogue of splice junction sequences";
RL Nucleic Acids Res. 10:459-472(1982).
RN [7]
RX MEDLINE; 92168112.
RA Adams M.D., Dubnick M., Kerlavage A.R., Moreno R., Kelley J.M.,
RA Utterback T.R., Nagle J.W., Fields C., Venter J.C.;
RT "Sequence identification of 2,375 human brain genes";
RL Nature 355:632-634(1992).
RN [8]
RX MEDLINE; 90295283.
RA Zhu Q.S., Heisterkamp N., Groffen J.;
RT "Characterization of the human ABL promoter regions";
RL Oncogene 5:885-891(1990).
RN [9]
RA Paskind M.;
RT "Personal communication";
RL Unpublished.
RN [10]
RA Tatusov R., Lipman D.J.;
RT "Using local similarities for pattern detection in
RT nucleotide/protein sequences";
RL Unpublished.
RN [11]
RX MEDLINE; 84106827.
RA Groffen J., Stephenson J.R., Heisterkamp N., De Klein A.,
RA Bartram C.R., Grosfeld G.;
RT "Philadelphia chromosomal breakpoints are clustered within a
RT limited region, bcr, on chromosome 22";
RL Cell 36:93-99(1984).
RN [12]
RP 1-35962
RA Chissee S.L.;
RT ;
RL Thesis (1994), University of Oklahoma
RN [13]
RP 1-35962
RA Roe B.A.;
RT ;
RL Submitted (08-MAR-1994) to the EMBL/GenBank/DBJ databases.
RL Bruce A. Roe, Dept. of Chem. and Biochem., University of Oklahoma,
RL 620 Parrington Oval, Rm 208, Norman, OK 73019, USA
CC NCBI gi: 514262
FH Key Location/Qualifiers
FH source 1..35962
FT /clone="Ca9-1"
FT /chromosomes="9"
FT /map="9q34"
FT /organism="Homo sapiens"
FT /tissue_type="lung carcinoma"
FT complement(448..783)
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT complement(478..(726.728))
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT

FT repeat_region 1012..(1327.1338)
FT /citation=[10]
FT /rpt_family="Alu"
FT complement(1455..1786)
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT (1907.1915)..2184
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT 2185..2227
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT complement(2619..2894)
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT 3454..3774
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT (4049.4097)..(4411.4433)
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT complement((5798.5804)..6146)
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT complement(6412..(6677.6691))
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT 6510..6820
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT complement((6980.6990)..(7289.7340))
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT <8604..8719
FT exon /number=1
FT /product="unknown"
FT join(8604..8719,10298..10399,12388..12433,14134..14199,
FT 15672..15740,16939..17115,17862..17990,18503..18583)
FT CDS /evidence=EXPERIMENTAL
FT /note="putative 8604 MET gene, 5' to the ABL gene; NCBI gi
FT :
FT 514266"
FT /citation=[6]
FT /citation=[7]
FT /codon_start=1
FT /product="unknown"
FT /db_xref="PID:g514266"
FT /translation="MEMRLPVARKELSERLGRDTRKHLVVPDITTTGTGEMRGCTY
FT GEEKLIASVAGSVERVKNKLCVKAALKTRYIGEVGDIIVGRI TERRRSAEDELAMRGFL
FT EGDLSAEVQAVFSDGAVSLHTRSIKYKLGQGVLVQVSPSLVKRQKTHFDLPCCAS
FT ILGNNGFIWYPTPEHKEEAGGFANLEPVSADREVISRINLCIISLVYQRMWLYD
FT SILYCYEASLPHQIKDILAKPEIMEEIVNMETRQRLLDEG"
FT complement(9524..9833)
FT repeat_region /citation=[10]
FT /rpt_family="Alu"
FT 10298..10399
FT exon /number=2
FT /product="unknown"
FT complement(10663..10902)
FT repeat_region /citation=[10]
FT

FT	repeat_region	/rpt_family="Alu"
FT		complement(10906..10991)
FT		/citations={10}
FT	exon	/rpt_family="Alu"
FT		12388..12433
FT		/number=3
FT	repeat_region	/product="unknown"
FT		complement(12599..12899)
FT		/citations={10}
FT	repeat_region	/rpt_family="Alu"
FT		13035..(13382..13389)
FT		/citations={10}
FT	repeat_region	/rpt_family="Alu"
FT		complement(13422..13425)..13719)
FT		/citations={10}
FT	repeat_region	/rpt_family="Alu"
FT		complement(13922..14172)
FT		/citations={10}
FT	exon	/rpt_family="Alu"
FT		14134..14199
FT		/number=4
FT	repeat_region	/product="unknown"
FT		(14674..14676)..(15029..15040)
FT		/citations={10}
FT	repeat_region	/rpt_family="Alu"
FT		complement(14984..15063)..(15359..15379))
FT		/citations={10}
FT	repeat_region	/rpt_family="Alu"
FT		complement(15133..15140)..15489)
FT		/citations={10}

Note: remainder of annotations omitted.

Query Match 15.3%; Score 238; DB 9; Length 35962;
Best Local Similarity 91.3%; Pred. No. 7.33e-104;
Matches 263; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

nb 15175 tttttttgagacggaattttcactctttattaccacagctggantgcaatggtggatctcg 15234

[illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

xy IAST GCTCAGATACCTCCGCGTCCAGGTTCAGGCATTCTCCGCTTAGGCCCTCCCTGGTA TGTG

Db 15295 gctgggattacaggcatgcccactatgcccgactaatttgtacttttagtagagacgg 15354

Qy 1091 GCTGGATTACAGGGCTGCCCACTATGCCCTGACTAATTTGTAGTTTAGTAGACGG 1150

Db 15355 ggtttctccatgttggtcaggctggtctagaactcctgacctcaggtgatcgacctgcct 15414

ov 1151 GGTTCCTCATGTTGGTCAGGCTGGTCTCAAACCTCCTGACCTCAGGTGATCTGCCCCGCT 1210

D6 15415 ccccccccct nct aggggt tgcggt gt acccccccacccacct ggg 15462

[illegible]

RESULT 5

DEFINITION Human ABL gene, exon 1b and intron 1b, and putative M8604 Met

protein (M0004 Met) gene, complete cos.
ACCESSION U07561

NID 951476Z

KEYWORDS	1 of 3
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryotae; Hyperchondria; Eukaryote crown group; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 35962)
AUTHORS	Chisose,S.L., Bodenteich,A., Wang,Y., Jian,L., Burian,D., Carriere,J.S., Freeman,A., Hamed,O.S., McLaury,H., Pan,H., Toth,S., Wang,Z., Zhang,G., Clifton,S., Groffen,J., Heisterkamp,N. and Roe,B.A.
TITLE	Human ABL gene region
JOURNAL	Unpublished (1994) Dept. of Chem. and Biochem., University of Oklahoma
REFERENCE	2 (sites)
AUTHORS	Groffen,J., Heisterkamp,N., Grosveld,F., Van de Ven,W. and Stephenson,J.R.
TITLE	Isolation of human oncogene sequences (v-fes homolog) from a cosmid library
JOURNAL	Science 216 (4550), 1136-1138 (1982)
MEDLINE	82199444
REFERENCE	3 (sites)
AUTHORS	Morris,C.M., Heisterkamp,N., Groffen,J. and Fitzgerald,P.H.
TITLE	Entire ABL gene is joined with 5'-BCR in some patients with Philadelphia-positive leukemia
JOURNAL	Blood 78 (4), 1078-1084 (1991)
MEDLINE	91329820
REFERENCE	4 (bases 27793 to 28338; 28792 to 29765)
AUTHORS	Stitelman,E., Lifshitz,B., Gale,R.P., Roe,B.A. and Canaani,E.
TITLE	Alternative splicing of RNAs transcribed from the human abl gene and from the bcr-abl fused gene
JOURNAL	Cell 47 (2), 277-284 (1986)
MEDLINE	87028219
REFERENCE	5 (bases 27108 to 29899)
AUTHORS	Bernards,A., Rubin,C.M., Westbrook,C.A., Paskind,M. and Baltimore,D.
TITLE	The first intron in the human c-abl gene is at least 200 kilobases long and is a target for translocations in chronic myelogenous leukemia
JOURNAL	Mol. Cell. Biol. 7 (9), 3231-3236 (1987)
MEDLINE	88038877
REFERENCE	6 (sites)
AUTHORS	Mount,S.M.
TITLE	A catalogue of splice junction sequences
JOURNAL	Nucleic Acids Res. 10 (2), 459-472 (1982)
MEDLINE	82150208
REFERENCE	7 (sites)
AUTHORS	Adams,M.D., Dubnick,M., Kerlavage,A.R., Moreno,R., Kelley,J.M., Uterback,T.R., Nagle,J.W., Fields,C. and Venter,J.C.
TITLE	Sequence identification of 2,375 human brain genes [see comments]
JOURNAL	Nature 355 (6361), 632-634 (1992)
MEDLINE	92168112
REFERENCE	8 (sites)
AUTHORS	Zhu,Q.S., Heisterkamp,N. and Groffen,J.
TITLE	Characterization of the human ABL promoter regions
JOURNAL	Oncogene 5 (6), 885-891 (1990)
MEDLINE	90295283
REFERENCE	9 (sites)
AUTHORS	Paskind,M.
TITLE	Personal communication
JOURNAL	Unpublished

```

REFERENCE 10 (sites)
AUTHORS Tatusov,R. and Lipman,D.J.
TITLE Using local similarities for pattern detection in
nucleotide/protein sequences
JOURNAL Unpublished
REFERENCE 11 (sites)
AUTHORS Groffen,J., Stephenson,J.R., Helsterkamp,N., de Klein,A.,
Bartram,C.R. and Grosveld,G.
TITLE Philadelphia chromosomal breakpoints are clustered within a limited
region, bcr, on chromosome 22
JOURNAL Cell 36 (1), 93-99 (1984)
MEDLINE 84106827
REFERENCE 12 (bases 1 to 35962)
AUTHORS Chisoe,S.L.
TITLE Sequence of the human abl and bcr genes
JOURNAL Thesis (1994) University of Oklahoma
REFERENCE 13 (bases 1 to 35962)
AUTHORS Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-1994) Bruce A. Roe, Dept. of Chem. and Biochem.,
University of Oklahoma, 620 Parrington Oval, Rm 208, Norman, OK
73019, USA
COMMENT NCBI gi: 514262
FEATURES
source
Location/Qualifiers
1..35962
/clone="Ca9-1"
/chromosomes="9"
/map="9q34"
/organism="Homo sapiens"
/tissue type="lung carcinoma"
/citation=[2]
/citation=[3]
/citation=[12]
complement(448..783)
/rpt family="Alu"
/citation=[10]
complement(478..726.728))
/rpt family="Alu"
1012..(1327.1338)
/citation=[10]
/rpt family="Alu"
/citation=[10]
complement(1455..1786)
/rpt family="Alu"
/citation=[10]
(1907.1915)..2184
/rpt family="Alu"
/citation=[10]
2185..2227
/rpt family="Alu"
/citation=[10]
complement(2619..2894)
/rpt family="Alu"
/citation=[10]
3454..3774
/rpt family="Alu"
/citation=[10]
(4049.4097)..(4411.4433)
/rpt family="Alu"
/citation=[10]
complement((5798.5804)..6146)
/rpt family="Alu"
/citation=[10]
complement(6412..(6677.6691))

```

```

/rpt family="Alu"
/citation=[10]
6510..6820
/rpt family="Alu"
/citation=[10]
complement((6980.6990)..(7289.7340))
/rpt family="Alu"
/citation=[10]
join(8604..8719,10298..10399,12388..12433,14134..14199,
15672..15740,16939..17115,17862..17990,18503..18583)
/note="putative 8604 MET gene, 5' to the ABL gene; NCBI
gi: 514266"
/codon_start=1
/evidence=experimental
/citation=[6]
/citation=[7]
/product="unknown"
/db_xref="PID:G514266"
/translation="MEMRLPVARKPLSERLGRDTKKHLVPGDTLTDTGFMRGHCTY
MGEKLIASVAGSVERNKLIQKALKTRYIGVGDIIVGRITERRRRAEDLNRGF
LQEGDLISAEQVAFSDCAVSLHTRSLKYGKLGQGLVQVPSLVKROKTHFDLPFG
ASVILGNNGFIWYPTPEHKEEAGGFANLEPVSLADREVISRINCIIISLVTRQWM
LYDTSILCYEASLPHQIKDKLPETMEEIIVMETRQRLJQEG"
<8604..8719
/number=1
/product="unknown"
complement(9524..9833)
/rpt family="Alu"
/citation=[10]
10298..10399
/number=2
/product="unknown"
complement(10663..10902)
/rpt family="Alu"
/citation=[10]
complement(10906..10991)
/rpt family="Alu"
/citation=[10]
12388..12433
/number=3
/product="unknown"
complement(12599..12899)
/rpt family="Alu"
/citation=[10]
13035..(13382.13389)
/rpt family="Alu"
/citation=[10]
complement((13422.13425)..13719)
/rpt family="Alu"
/citation=[10]
complement(13922..14172)
/rpt family="Alu"
/citation=[10]
14134..14199
/number=4
/product="unknown"
(14674.14676)..(15029.15040)
/rpt family="Alu"
/citation=[10]
complement((14984.15063)..(15359.15379))
/rpt family="Alu"
/citation=[10]
complement((15133.15140)..15489)
/rpt family="Alu"

```

repeat_region

repeat_region

CDS

exon

repeat_region

exon

repeat_region

repeat_region

exon

repeat_region

repeat_region

repeat_region

repeat_region

exon

repeat_region

repeat_region

repeat_region

repeat_region	/tissue_type="lung carcinoma" complement(448..783) /citation={10} /rpt_family="Alu"
repeat_region	complement(478..(726.728)) /citation={10} /rpt_family="Alu"
repeat_region	1012..(1327.1338) /citation={10} /rpt_family="Alu"
repeat_region	complement(1455..1786) /citation={10} /rpt_family="Alu"
repeat_region	(1907.1915)..2184 /citation={10} /rpt_family="Alu"
repeat_region	2185..2227 /citation={10} /rpt_family="Alu"
repeat_region	complement(2619..2894) /citation={10} /rpt_family="Alu"
repeat_region	3454..3774 /citation={10} /rpt_family="Alu"
repeat_region	(4049.4097)..(4411.4433) /citation={10} /rpt_family="Alu"
repeat_region	complement(45798.5804)..6146 /citation={10} /rpt_family="Alu"
repeat_region	complement(6412..(6677.6691)) /citation={10} /rpt_family="Alu"
repeat_region	6510..6820 /citation={10} /rpt_family="Alu"
repeat_region	complement((6980.6990)..(7289.7340)) /citation={10} /rpt_family="Alu"
exon	<8604..8719 /number=1 /product="unknown"
CDS	join(8604..8719,10298..10399,12388..12433,14134..14199, 15672..15740,16399..17115,17862..17990,18503..18583) /evidence=experimental /note="putative 8604 MET gene, 5' to the ABL gene; NCBI gi: 514266" /citation={6} /citation={7} /codon_start=1 /product="unknown" /db_xref="PID:g514266" /translation="MEMRLPVARRK LSELRGRDTKKHLLVPGDITTTDTGFMRGHGY MGEKLIASVAGSVEVRNKKLCVKAALKTRYIGEVGDIIVGRITERRBSAEDELARMGF LQEGDILSAEQVAFSDGASVLSHRSKYKLGCGVLVQVSFLVKQKTHFDLPDGG ASVILGNNGFTWYPTPEKKEEAGGFIANLEPYSIADREVISLRNCTIISLTQRMW LYDTSILCYCEASLPKHQDKLPEIMEEIVMETRQLRLEQSG"
repeat_region	complement(9524..9933) /citation={10} /rpt_family="Alu"
exon	10298..10399 /number=2 /product="unknown"

```

repeat_region complement (10663..10902)
/citation={10}
/rpt_family="Alu"
repeat_region complement (10906..10991)
/citation={10}
/rpt_family="Alu"
exon 12388..12433
/number=3
/product="unknown"
repeat_region complement (12599..12899)
/citation={10}
/rpt_family="Alu"
repeat_region 13035..(13382..13389)
/citation={10}
/rpt_family="Alu"
repeat_region complement ((13422..13425)..13719)
/citation={10}
/rpt_family="Alu"
repeat_region complement (13922..14172)
/citation={10}
/rpt_family="Alu"
exon 14134..14199
/number=4
/product="unknown"
repeat_region (14674..14676)..(15029..15040)
/citation={10}
/rpt_family="Alu"
repeat_region complement ((14984..15063)..(15359..15379))
/citation={10}
/rpt_family="Alu"
repeat_region complement ((15133..15140)..15489)
/citation={10}
/rpt_family="Alu"
repeat_region 15367..15603
/citation={10}
/rpt_family="Alu"
repeat_region complement (15562..15812)
/partial
/citation={10}
/rpt_family="Alu"
...
Note: remainder of annotations omitted.

Query Match 15.3%; Score 238; DB 86; Length 35962;
Best Local Similarity 91.3%; Pred. No. 7.33e-104;
Matches 263; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Db 15175 ttttttgacggagtttcaactcttgttccacggtggagtgcaatgggtgggatctcg 15234
|||||
Qy 971 TTTTITGATGAGAGTTTCGGCTTTTGTGCCAGCGTGAGTGCAATGGCGGATCTTG 1030
|||||

Db 15235 gctcactgcaacctctgcctccgggttcaagcgattctcctgcctcagctcactgagta 15294
|||||
Qy 1031 GCTCACCATAACCTCCGCTCCAGGTTCAAGCAATTCCTCGCTTAGCCTCCTCGAGTA 1090
|||||

Db 15295 gctgggattacaggatgcgcacactatgcccgactaatittgtacttttagtagacagcg 15354
|||||
Qy 1091 GCTGGGATTACAGGCGTGCGCCACTATGCCTGACTAATTTGTAGTTTTAGTAGACAGCG 1150
|||||

Db 15355 ggtttctccatggttggtcaggtgggtctagaactcgtgcactcaggtgatcgactgcct 15414
|||||
Qy 1151 GGTTCCTCATCTGGTCAGGCTGGTCTCAAACTCCTGCACCTCAGGTGATCTGCCGCT 1210
|||||

Db 15415 caagctcccaaaagtctgggattacaggtgtgagccaccgcacactggc 15462

```

```

QY 1211 CAGCCTCCCAAGTCTGGAATTACAGCGGTGAGCCACCGCCCTGGC 1258
|||||
RESULT 7
LOCUS HSATPCP1 9457 bp DNA PRI 19-JUL-1993
DEFINITION H.sapiens gene for mitochondrial ATP synthase c subunit (P1 form).
ACCESSION X69907
NID 938429
KEYWORDS ATP synthase; ATP synthase c subunit.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 9457)
AUTHORS Walker, J.E.
TITLE Direct Submission
JOURNAL Submitted (29-DEC-1992) to the EMBL/GenBank/DBJ databases. J.E.
Walker, M.R.C. Lab. of Molecular Biology, Hills Road, Cambridge CB2
2QH., UK
REFERENCE 2 (bases 1 to 9457)
AUTHORS Dyer, M.R. and Walker, J.E.
TITLE Sequences of members of the human gene family for the c subunit of
mitochondrial ATP synthase
JOURNAL Biochem. J. 293 (Pt 1), 51-64 (1993)
MEDLINE 93319529
COMMENT NCBI gi: 38429
FEATURES
source
1..9457
/organism="Homo sapiens"
/cell_type="T-cell"
/cell_line="AT5B1"
/clone_lib="lambda 2001"
1..224
/rept_family="Alu"
repeat_unit 273..442
/rept_family="Alu"
repeat_unit 606..904
/rept_family="Alu"
repeat_unit 1449..1768
/rept_family="Alu"
repeat_unit 1769..2111
/rept_family="Alu"
repeat_unit 2627..2945
/rept_family="Alu"
repeat_unit 2646..3233
/rept_family="Alu"
GC signal 4041..4046
repeat_unit 4100..4792
/rept_family="Alu"
CAAT signal 4538..4542
CAAT signal 4641..4645
CAAT signal 4671..4675
CAAT signal 4697..4701
TATA signal 4708..4713
mRNA join(4733..4842,5341..5388,6304..6381,7088..7266,
7587..7802)
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
prim_transcript 4733..7802
/gene="P1 gene for c subunit of human mitochondrial ATP

```

```

synthase"
4733..4842
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=1
4843..5340
/number=1
5002..5007
5279..5280
5341..5388
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=2
join(5350..5388,6304..6381,7088..7266,7587..7701)
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/note="NCBI gi: 38430"
/codon_start=1
/db_xref="PID:q38430"
/translation="MQTAGALFISPALRCRGLIRVVSASFINSVPNSKQPSYSN
FPLQVAREFQTSVSRDIDTAKEIGAGAAATVGAGGAGICTVFGSLIIGYARNPS
IKQQLFYAILGELSEAMGLFCIMVAFILFAM"
5389..6303
/number=2
6304..6381
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=3
6382..7087
/number=3
6418..6699
/rept_family="Alu"
7088..7266
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=4
join(7154..7266,7587..7698)
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
7267..7586
/number=4
7587..7802
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=5
7780..7785
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
8879..9176
/rept_family="Alu"
BASE COUNT 2505 a 2208 c 2318 g 2426 t
ORIGIN
Query Match 15.2%; Score 237; DB 49; Length 9457;
Best Local Similarity 90.7%; Pred. No. 2.5le-103;
Matches 264; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
Db 8878 agccagcgatggtgcacgcctgtaatcccagcacttggaggagtgagcgcgga 8937
|||||
Cp 1259 AGCCAGCGGTGGTGGCTCAGCGCTGTAATTCAGACACTTTGGGAGGCTGAGCGGCGGAGA 1200
|||||
Db 8938 tcacctgaggtcaggagttcgaccagcctggaacaacatggagaacccagctctact 8997
|||||
Cp 1199 TCACCTGAGGTGAGGAGTTTCAGACACCGCTGACCAACATGAGAACCCGCTCTACT 1140
|||||

```

```

NCBI gi: 34286
FEATURES             Location/Qualifiers
     source            1..6901
                        /organism="Homo sapiens"
                        /clone_lib="(lambda)charon30"
                        CAAT_signal 709..716
                        TATA_signal 786..790
                        prim_transcript 809..5052
                        mRNA          4458..5052
                        exon          809..990
                        /number=1
                        CDS           join(837..990,1724..1880,1960..2075,2170..2265,2349..2573,
                                     4458..5032)
                                     /note="NCBI gi: 34287"
                                     /codon_start=1
                                     /product="lecithin-cholesterol acyltransferase (LCAT)"
                                     /db_xref="PID:g34287"
                                     /trnaelation="MGCPGSGQWVTLGLLPPAAPFWLLNLFPPHTPKAELSN
                                     TRVTVNLSPGCGNQLEAKLDQDVVNMCMKRYEDFTTILDLNMFPLGVDCMDIN
                                     TRVTVNRSGLVSNAPGVQIRVPGKTTSEVYLDSSKLAGYLHTLVQNLVNNVGTRD
                                     ETRAPVYDRLLEPQQEEYKELAGLVEEMHAAVKGKPVFLIHSIAGLHLLYLPQW
                                     EQWAKDFIDGFI SLGAPGGGSKPMVLVLSAGDNQGITPMSISKKEQRIITTSPNMW
                                     FFSMAWPEDHVFISPTSNYTGDFORFFADLHFEQGWYMLQSRDLIAGLAPGAPVGE
                                     VYCLYGVLPTPTYIDYHGFFYTDPPGVLYEDGDDTVAIRSTELCGMGQGPQCPVH
                                     LLP LHGILQHNNVSNLTLEHINALIGAYTQGGPASPATSPPEPPPE"
                                     intron 991..1723

```

Query Match	14.9%;	Score 232;	DB 51;	Length 6901;
Best Local Similarity	88.9%;	Pred. No. 1.19e-100;		
Matches 265;	Conservative	0;	Mismatches 33;	Indels 0; Gaps 0;
Db	6351	agcctgtgctgtgcctcatcctgtataccccagcaactctgggagcgcgagctggggcgga	6410	
Cp	1259	AGCCAGCGCTGGTGGCTCAGCGCTGTAATTCGACCACTTTGGGAGGCTGAGGGGGGCGA	1200	
Db	6411	tcacttgaggtcgaaagttdgagaccaccctgcaccaactggagaaacccgcctctctact	6470	
Cp	1199	TCACCTGAGTCTGAGGATTTGAGCACGCGCTGCACACATGAGAAACCCGCTCTCTACT	1140	
Db	6471	aaaaatcacaaattagccggcgatggtggcgcatgcctgtaatcccagctactcgggag	6530	
Cp	1139	AAAACTCAGAAAATTACTCAGGCATAGTGGCGCAGCCTCTAATCCAGCTACTCAGGAGG	1080	
Db	6531	ctgaggcaggagaattgtttaacctgggaggtggagcgttgagcgttgagccaaagatcacac	6590	
Cp	1079	CTAAGSGAGGAAATTGCTTGAACTTGGGAGGCGGAGGTATGTGTCAGCCAAAGTCGGC	1020	
Db	6591	tattgcaactccagcctggggcacaagaagcaaaactcgcgtctcaaaaaaaaaatttat	6648	
Cp	1019	CATTGCATCCAGCTTGGGCAACAAAGGCGAACTCATCTCAAAAAAAGACGTAT	962	

RESULT	9	HSTMP04	6905 bp	DNA	PRI	02-JUL-1995
LOCUS						

DEFINITION Human thymopoietin (TMPO) gene, exons 4 and 5, and complete cds for thymopoietin alpha.

ACCESSION U18270

NID 9885679

KEYWORDS

SEGMENT 4 of 6

SOURCE human.

ORGANISM

Eukaryotes; mitochondrial eukaryotes; eukaryote crown group; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Harris, C.A., Andryuk, P.J., Cline, S.W., Siekierka, J.J. and Goldstein, G.

TITLE

Structure and mapping of the human thymopoietin (TMPO) gene and relationship of TMPO beta with rat lamin-associated polypeptide 2

JOURNAL

Unpublished

AUTHORS

Harris, C.A.

TITLE

Direct Submission

JOURNAL

Institute, Route 22 East, Annandale, NJ 08801-0999, USA

NCBI gi: 885679

COMMENT

FEATURES

source

1..6905

/clone="lambda SHC-1"

/clone lib="Placenta genomic library in lambda FIX II;

Stratagene Catalog no.946203"

/chromosome="12"

/organism="Homo sapiens"

/map="12q22; 64% (% distance from centromere to telomere)"

join(U18266:2201...2479,U18267:120..246,U18268:130..288,

39..1558)

/gene="TMPO"

/note="NCBI gi: 885683"

/codon_start=1

/product="thymopoietin alpha"

/db_xref="PID:9885683"

/translation="MFELEDPSVLTQKLSLVANNVLPAGEQRKDVYVLIYQH

LTAARNRPPIAGTNSKGPDPFSSDEERPTVLGSGAAAGRAAVGRKATKTKDP

ROEKDLDVTELTNEDLDQIVKYGVNPGP IVGTRKLYEKKLIREQGTESRST

PLPTISSAENTRQNSDRYSNDEGKKHKKVSTRDIVPSELGTFSGGGF

FQGISFPEIETRPPLGSLTEQAARKVHTSKGDLPREPLVATNLPGRGLQKLASERNL

FTSCSSSHDRCLKSSSSSQPHSAMLVSTAAASPLIKETTTGYKDIVENTIGREK

SGIQLPERSHISDQPLSSKKKALESSSOLISPLAARDVNSLVUGGVGS

LPFTSNMPLDVENIQKIDQSKQTEFELSPRPKVRPSKSEVERDSGSVAFQN

IPGSELMSFPAKTVVSHKLTTLGLVAKOSQHDKIDASELSFPFHEIILKVIEEQQ

VDRQPLSLACKTPVSSREATQILSPKVDDEILGFISEATPLGGIQAATESCNQQLD

LALCRAYEAAASAIQIATHTAFVAKAMQADISEAQILSSDPSTHQALGILSKTYDA

ASYICEAAFDVKMAAHTMGNAVTRRYLWIKOCKINLASKNKLASTPFGKGLFGE

VCKVTKRGNKH"

39..one-of(1956,2843)

/gene="TMPO"

/note="TMPO alpha-specific exon"

/number=4

4691..4788

/gene="TMPO"

/number=5

2028 a 1280 c 1359 g 2238 t

BASE COUNT

ORIGIN

Query Match

14.9%; Score 232; DB 52; Length 6905;

Best Local Similarity 90.0%; Pred. No. 1.19e-100;
Matches 261; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Db 6172 gccgggagtgagtcacgctgtatccagcaactttggggagcgagggcgagat 6231

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Cp 1258 GCCAGGCGTGGTGGCTCAGCGCTGTAATTCAGCACTTTGGGAGGCTGAGGGCGGAGAT 1199

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 6232 cactgaggtccgaggttcgagcgagcctgaccaacatggaggaacccctgtctacta 6291

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Cp 1198 CACTGAGGTGAGGTTTGAGCAGCGCTGACCAATGAGGAGAACCCCGTCTCTACTA 1139

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 6292 aaatacaaaattagcagcgatggtgcatcgctgtaactcctagctactccggagggc 6351

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Cp 1138 AAATACAAAATTAGTCAGGATAGTGGCGACCGCTGTAATCCCACTACTCAGGAGGC 1079

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 6352 tgaggcaggaatcgcttgaaaccttgaggcgaggtgtgtgagcgagatcgggc 6411

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Cp 1078 TAAAGCAGGAGAAATTGCTGAACCTGGGAGCGGAGGTTATGCTGAGCCAGATCGGGC 1019

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 6412 attgcaacttagctgggcaacagaggaactctgtctcaaaaaataa 6461

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Cp 1018 ATTGCACTCAGCGCTGGCAACAAAGCGAACTCCATCTCAAAAAAAA 969

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10

LOCUS HSPROSCHY 13863 bp DNA PRI 11-NOV-1993

DEFINITION H.sapiens genes for proteasome-like subunit (MECL-1),

chymotrypsin-like protease (CTRL-1) and protein serine kinase

(PSK-HI) last exon.

ACCESSION X71874

NID 9406226

KEYWORDS CpG island; gene cluster; proteasome subunit; serine protease;

serine protein kinase.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa/Eumycota group;

Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;

Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;

Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;

Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 13863)

AUTHORS Larsen, F.

TITLE Direct Submission

JOURNAL Submitted (10-MAY-1993) to the EMBL/GenBank/DBJ databases. F.

Larsen, Biotechnology Centre of Oslo, University of Oslo, PO Box

1125 Blindern, N0317 Oslo, NORWAY

REFERENCE 2 (bases 1 to 13863)

AUTHORS Larsen, F., Solheim, J., Kristensen, T., Kolsto, A.B. and Prydz, H.

TITLE A tight cluster of five unrelated human genes on chromosome 16q22.1

JOURNAL Hum. Mol. Genet. 2 (10), 1589-1595 (1993)

COMMENT 94093544

NCBI gi: 406226

FEATURES

source

1..13863

/organism="Homo sapiens"

/clone.lib="cosmid COS202"

/clone="cosODIN"

/map="16q22.1"

/cell_type="leukocytes"

repeat_unit 123..423

/note="Alu repeat"

repeat_unit 513..805

/note="Alu repeat"

RBS 566..571

```
repeat_unit /note="putative Spl"
820..1131
/note="Alu repeat"
repeat_unit 1134..1296
/note="Alu repeat"
repeat_unit 1350..1482
/note="Alu repeat"
RBS 1576..1581
/note="putative Spl"
2251..3970
/note="CpG island"
TATA_signal 2267..2272
RBS 2308..2313
/note="putative Spl"
<2310..2400
/number=1
join(2345..2400,2609..2696,2782..2879,2991..3131,
3395..3510,3615..3673,4145..4296,4422..4533)
/note="NCBI gi: 406227"
/codon_start=1
/product="protease-like subunit MECL-1"
/db_xref="PID:g406227"
/translation="MIRPALEPRGGFSCFNQNASLERVLPGLKVPHARKTGTIAG
LVFGDVLGAOTRATNDVVADKCEKIHFIAPKIYCGGAGVADAEMTTTRVASKM
ELHALSTGRFPAVTVRIIRQTLFRQGVHGASLIVGGVDLTGQLYGVHPHGSYSR
LPFTALGSGODAAALAVLEDLRFQPNMTLEAAQGLLVEAYTAGILGDLGGGNDACVIT
KTGAKLRLTSSPTPEVKRSGRYHEVPQTAVLTQVXP LITLVEETVQAMEVE"
2345..2400
/misc_feature /label=test
intron 2401..2608
/number=1
RBS 2419..2424
/note="putative Spl"
exon 2609..2696
/number=2
intron 2697..2781
/number=2
exon 2782..2879
/number=3
intron 2880..2990
/number=3
exon 2991..3131
/number=4
intron 3132..3394
/number=4
RBS 3208..3212
/note="putative Spl"
RBS 3376..3381
/note="putative Spl"
exon 3395..3510
/number=5
RBS 3428..3433
/note="putative Spl"
intron 3511..3614
/number=5
exon 3615..3673
/number=6
intron 3674..4144
/number=6
exon 4145..4296
/number=7
intron 4297..4421
/number=7
polyA_signal 4363..4568
```

```
exon 4422..>4568
/number=8
misc_feature 4422..4533
/label=test2
repeat_unit 5560..6000
/note="Alu repeat"
RBS 5942..5948
/note="putative Spl"
6037..6326
/note="Alu repeat"
TATA_signal 7202..7207
prim_transcript 7232..9523
/note="C"
prim_transcript 7232..9484
/note="G"
prim_transcript 7232..9210
/note="A3"
prim_transcript 7232..9220
/note="A2"
prim_transcript 7232..9216
/note="A1"
exon 7232..7293
/number=1
misc_feature 7242..7293
/label=st
join(7242..7293,7893..7996,8095..8174,8287..8368,
8506..8686,8774..8907,8998..9159)
/note="NCBI gi: 406228"
/codon_start=1
/product="chymotrypsin-like protease CTRL-1"
/db_xref="PID:g406228"
/translation="MLLSLTSLVLGSSWGGGIPAIKPAISFSQRIINGENAVIGS
WPQWSIQDSSGFHFCGSLISQSWVTAAHCNVPGRHFVILGEYDRSSNAEP LQVL
SVSRATHPSSWSTTMANDVTILKLA SPAQYTTTRISPVCLASNEALTEGTCVTTCG
GRLSGNGVTPAHLQQVALPLVTVNCQKQYWGSSITDSMICAGGASSCGSGGGL
VQKGTWVLIGIVSWGTCNCRAPAVTVTRVSKFTWVQVIAYN"
join(7242..7293,7893..7894)
sig_peptide 7242..7293
exon /note="signal peptide"
/number=1
intron 7294..7892
/note="signal peptide"
/number=1
exon 7893..7996
/number=2
exon 7893..7894
/note="signal peptide"
/number=2
misc_feature 7895..7939
/note="activation peptide"
misc_feature 7940..7996
/label=stl
join(7940..7996,8095..8174,8287..8368,8506..8686,
8774..8907,8998..9159)
/mat_peptide /product="chymotrypsin-like protease CTRL-1"
7997..8094
/number=2
intron 8095..8174
exon 8175..8286
intron 8287..8368
exon 8369..8505
intron
```

```

exon          /number=4
              8506..8686
intron        /number=5
              8687..8773
exon          /number=5
              8774..8907
intron        /number=6
              8908..8997
              /number=6
              8998..9159
misc_feature  /label=e1
              8998..9159
misc_feature  /label=e
              8998..9216
exon          /note="number 7A1"
              8998..9523
exon          /note="number 7C"
              8998..9484
exon          /note="number 7B"
              8998..9210
exon          /note="number 7A3"
              8998..9220
exon          /note="number 7A2"
              9190..9195
polyA_signal  /note="number A"
              complement(9415..11763)
exon          /number=2
              9415..9523
prim_transcript complement(9415..13863)
misc_feature  /note="overlapping transcription unite"
              9462..9467
polyA_signal  /note="number B"
              9507..9512
polyA_signal
...
Note: remainder of annotations omitted.

```

```

Query Match      14.9%; Score 232; DB 52; Length 13863;
Best Local Similarity 88.9%; Pred. No. 1.19e-100;
Matches 265; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Db 513 agcctggtgcgtggtcctgctgtaatccagcactctggaggccgagtggtggcga 572
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1259 AGCAGGCGGTGGTGGCTCAGCCCTGTAATCCAGCCTTTGGGAGGCTGAGCGGCGAGA 1200

Db 573 tcacttgaggtcagaagtgtgagaccacctgaccacaatggagaacccccgtctctact 532
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1199 TCACCTGAGGTGAGGAGTTTGAGCAGCGCTGACCAACATGACGAACCCCTCTCTACT 1140

Db 633 aaaaatacaaaattagcggcgatggtggcgtgctgtaatccagctactcggagg 692
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1139 AAMACTACAAATTAGTCAGGCATAGTGGCGCAGCGCTGTAATCCAGCTACTCAGGAGG 1080

Db 693 ctgagcgagagaattgttgaacctggaggtggacgttgcgtgagcgaagatcacac 752
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1079 CTAAAGCAGGAAATTTGCTTCAACCTGGGAGGCGGAGTTATGCTGAGCAAGATCGCCG 1020

Db 753 tattgcactccagctggggcaagaagacaactccgtctcaaaaaaaatttat 810
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1019 CATTCGACTCCAGCTGGGCAACAAGACGAACCTCATCTCAAAAAAAGAGCTAT 962

RESULT 11
ID HSL69F7B standard; DNA; PRI; 16851 BP.
AC 269650;

```

```

DT 22-FEB-1996 (Rel. 46, Created)
DE 22-FEB-1996 (Rel. 46, Last updated, Version 1)
DE Human DNA sequence from cosmid L69F7B, Huntington's Disease Region,
DE chromosome 4p16.3 contains Huntington Disease (HD) gene.
KW 4p16.3; Huntington Disease (HD) gene.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Euthera; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-16851
RA Mungall A, Odell C.;
RT ;
RL Submitted (16-FEB-1996) to the EMBL/GenBank/DBJ databases.
RL Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, UK. E-mail
RL enquires: humquery@sanger.ac.uk
CC IMPORTANT:
CC This sequence is not the entire insert of clone L69F7B. It may be
CC shorter because we only sequence overlapping sections once, or
CC longer because we arrange for a small overlap between neighbouring
CC submissions.
CC This sequence has been finished according to sequence map criteria
CC as follows. An attempt is made to resolve all sequencing problems,
CC such as compressions and repeats, but not necessarily within known
CC annotated human repeat sequence elements (e.g. Alu). Where the
CC sequence is ambiguous, there is an annotation using the "unsure"
CC feature key.
CC The sequence from clone L161A8 has been finished in more than one
CC contig.
CC This sequence (L161A8B) is separated from the preceeding one
CC (L161A8A) by
CC a gap of 1100bp.
CC The true left end of clone L161A8 is at 16748.
CC L69F7 is from cosmid library LA04NC01 constructed at the
CC Human Genome Center, Los Alamos National Laboratory, NM 87545
CC under the auspices of the U.S. Department of Energy.
CC The library was constructed using flow-sorted human chromosome 4
CC from a Hamster-Human hybrid cell line (UV20HL21-27) containing
CC human fchromosomes 4, 8 and 21.
CC VECTOR: sCos1
CC L69F7 is contained in a clone contig spanning ~2Mb which is
CC described in Baxendale et al, Nature Genetics 4 (1993) 181-
CC 186. See also Myers et al, Cytogenet Cell Genet. 66 (1994)
CC 218-230.
FH Key Location/Qualifiers
FH source 1..16851
FH /organism="Homo sapiens"
FH /chromosome="4"
FH /map="4p16.3"
FH /clone="L69F7B"
FH /clone_lib="LA04NC01"
FH /cell_line="UV20HL21-27"
FH CDS complement(join(<1187..1313,7404..7606,8663..8815,10383..
FH 10459,12716..12836,14946..>15092))
FH /product="Huntington Disease (HD) gene"
FH /partial
FH /codon_start=2
FH /db_xref="pID:e"
FH repeat_region 1316..1353
FH /note="19 copies of 2 mer 84 % conserved"
FH repeat_region 4203..4240
FH /note="19 copies of 2 mer 84 % conserved"
FH repeat_region 2282..2392
FH /note="MSTC element fragment"

```

FT repeat_region 2313..2366
/note="MLT1A element fragment"
FT repeat_region 2436..2572
/note="MLT1A element fragment"
FT repeat_region 2497..2571
/note="MLT1C element fragment"
FT repeat_region 3026..3064
/note="L1 element fragment"
FT repeat_region 3109..3155
/note="L1 element fragment"
FT repeat_region 3831..4026
/note="MER1B element fragment"
FT repeat_region 4022..4128
/note="MSTA element fragment"
FT repeat_region 4241..4452
/note="MSTA element fragment"
FT repeat_region 4450..4532
/note="MER1B element fragment"
FT repeat_region 5242..5435
/note="MER3 element fragment"
FT repeat_region 81..374
/note="Alu repeat: matches 1..308 of consensus"
FT repeat_region 2666..2889
/note="Alu repeat: matches 1..236 of consensus"
FT repeat_region 3218..3504
/note="Alu repeat: matches 1..308 of consensus"
FT repeat_region 3505..3674
/note="Alu repeat: matches 122..308 of consensus"
FT repeat_region 6004..6306
/note="Alu repeat: matches 1..308 of consensus"
FT repeat_region 10993..11279
/note="Alu repeat: matches 1..308 of consensus"
FT repeat_region 11703..11994
/note="Alu repeat: matches 1..308 of consensus"
FT repeat_region 16105..16396
/note="Alu repeat: matches 1..308 of consensus"
FT repeat_region 16691..16851
/note="Alu repeat: matches 1..177 of consensus"
FT repeat_region 15550..15627
/note="Alu repeat: matches 82..1 of consensus"
FT repeat_region 15332..15549
/note="Alu repeat: matches 308..75 of consensus"
FT repeat_region 13831..14121
/note="Alu repeat: matches 308..1 of consensus"
FT repeat_region 12159..12446
/note="Alu repeat: matches 308..1 of consensus"
FT repeat_region 7098..7241
/note="Alu repeat: matches 158..1 of consensus"
FT repeat_region 6804..7095
/note="Alu repeat: matches 308..1 of consensus"
FT repeat_region 6645..6803
/note="Alu repeat: matches 308..149 of consensus"
FT repeat_region 4563..4856
/note="Alu repeat: matches 308..1 of consensus"

FT repeat_region 1682..1966
/note="Alu repeat: matches 308..1 of consensus"
FT repeat_region 16851 BP; 5048 A; 3835 C; 3519 G; 4449 T; 0 other;
Matches 261; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Query Match 14.9%; Score 232; DB 10; Length 16851;
Best Local Similarity 90.0%; Pred. No. 1.19e-100;
Matches 261; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Db 11704 gccaggcatggtgtcacacctgtatccagcactttggggggtcgaggaggcgat 11763
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Cp 1258 GCCAGCGGTGGCTCAGCGCTGTATTCCAGCACTTTGGAGGGCTGAGCGGCGAGAT 1199
Db 11764 cacttgaggttggagtttgagaccagctgacaaacatggagaaacccctgtctacta 11823
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Cp 1198 CACCTCAGGTCAGGAGCTTGAGACCGCTGACCAACATGGAGAACCCCGTCTACTA 1139
Db 11824 aaatataaaattagccgggcatggtgtgcctgtaatccagctactcaggagc 11883
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Cp 1138 AAATACAAAATTAGTCAGGCATAGTGGCGACGCTGTAAATCCAGCTACTCAGGAGC 1079
Db 11894 tggaggcaggagaattgttgtaacccggaggtggtgaggttggtagtgagatcatgcc 11943
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Cp 1078 TAGGCGAGCAATTGCTTGACCTGGGAGCGGAGGTTATGCTGAGCCCAAGATCGGCGC 1019
Db 11944 attgcactccagctgggcaagaagcaaaactctgtctcaaaaaaa 11993
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Cp 1018 ATTGCATCCAGCGCTGGCGACCAACAAAAGCGAAATCCATCTCAAAAAAAA 969
RESULT 12
ID HSF1AG6PD standard; DNA; PRI; 219447 BP.
AC 144140;
DT 29-FEB-1996 (Rel. 47, Created)
DE Homo sapiens chromosome X region from filamin (FLN) gene to
DE glucose-6-phosphate dehydrogenase (G6PD) gene, complete cds's.
KW 1A gene; 2.19 gene; ABP-280 gene; actin-binding protein;
KW DNase I-like protein gene; DNLI1L gene; EMD gene; emerin;
KW emery-dreifuss syndrome; filamin; FIN gene; G4.5 gene; G4.8 gene;
KW G6PD gene; gdi gene; Gdx gene; glucose-6-phosphate dehydrogenase;
KW P3 gene; QW gene; STA gene; XAP-1 gene; XAP-2 gene; XAP-4 gene;
KW XAP-5 gene; XAP-7 gene.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-219447
RA Chen E.Y., Zollo M., Mazzarella R.A., Ciccodicola A., Chen C.N.,
RA Zuo L., Helmer C., Burrough F.W., Ripetto M., Schlessinger D.,
RA D'Urso M.;
RT "Thirteen known and six candidate genes in 219.4kb of high GC DNA
RT between the human RCP/GCP and G6PD loci";
RL Unpublished.
CC Submitted by: Ellison Chen, Advanced Center for Genetic Technology,
CC Applied Biosystems Division of Perlin Elmer Corp., 850 Lincoln
CC Center Drive, Foster City, CA 94044 USA and David Schlessinger,
CC Department of Molecular Microbiology and Center for Genetics in
CC Medicine Washington University School of Medicine, St. Louis MO
CC 63110 USA e-mail: ellison@genseq.apltdbio.com and
CC david@genetics.wustl.edu Note: Gene predictions were accomplished
CC with runs of Grail versions 1.1 and 1.2, coupled with fasta and
CC blastx comparisons to genbank & non-redundant peptide libraries.

Aug 27 08:17

US-08-231-565A-1.rge

31

CC Repeat analysis was accomplished via censor. NCBI gi: 1203968

Key	Location/Qualifiers
source	1..219447 /organism="Homo sapiens" /sequenced_mol="DNA" 4..223 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sp or Alu-Sq" 226..503 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 641..931 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 1023..1313 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(1446..1536) /evidence=EXPERIMENTAL /map="X"
repeat_region	/rpt_family="Alu-J" 3095..3426 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="CpG Island" complement(3663..3952) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sb0 or Alu-Sb1" 4912..5046 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 5047..5333 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sb0 or Alu-Sb1" 5346..5628 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 5632..5805 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sq" 6363..17698 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="11Kb repeat1" 6369..6633 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(7121..7409) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sp" complement(7358..7432) /evidence=EXPERIMENTAL

Aug 27 08:17

US-08-231-565A-1.rge

32

repeat_unit	/map="X" /rpt_family="Alu-Sb0 or Alu-Sb1" complement(7426..7701) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J" 8306..8592 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(8863..9029) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" complement(9031..9317) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sq" complement(9318..9442) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 9486..9773 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J" 9996..10132 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sp" 10159..10255 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sq" 10271..10367 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J" complement(10455..10739) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(10740..10860) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" complement(10866..11156) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" 11223..11501 /evidence=EXPERIMENTAL /map="X"
repeat_region	/rpt_family="Alu-J" 11638..13564 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="CpG Island" complement(11884..12163) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sb0 or Alu-Sb1" complement(12195..12482) /evidence=EXPERIMENTAL

Qy	1209	CTCAGCGCTCCCAAGTCTGGCAATTACGCGTGAGCCACCAACGCGCTGGCT	1259
RESULT	13		
ID	HSFUNG6PD	standard; DNA; PRI; 219447 BP.	
AC	L44140;		
DT	29-FEB-1996	(Rel. 47, Created)	
DT	29-FEB-1996	(Rel. 47, Last updated, Version 2)	
DE	Homo sapiens chromosome X region from filamin (FLN) gene to		
DE	glucose-6-phosphate dehydrogenase (G6PD) gene, complete cds' s.		
DE	1A gene; 2.19 gene; ABP-280 gene; actin-binding protein;		
KW	DNASE I-like protein gene; DNLL1 gene; EMD gene; emerlin;		
KW	emery-dreifuss syndrome; filamin; FLN gene; G4.5 gene; G4.8 gene;		
KW	G6PD gene; gdi gene; Gdx gene; glucose-6-phosphate dehydrogenase;		
KW	P3 gene; QM gene; STA gene; XAP-1 gene; XAP-2 gene; XAP-4 gene;		
KW	XAP-5 gene; XAP-7 gene.		
OS	Homo sapiens (human)		
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;		
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homiidae.		
RN	[1]		
RP	1-219447		
RA	Chen E.Y., Zollo M., Mazzarella R.A., Ciccodicola A., Chen C.N.,		
RA	Zuo L., Helner C., Burrough F.W., Ripetto M., Schlessinger D.,		
RA	D'Urso M.;		
RT	"Thirteen known and six candidate genes in 219.4kb of high GC DNA		
RT	between the human RCP/GCP and G6PD loci";		
RL	Unpublished.		
CC	Submitted by: Ellson Chen, Advanced Center for Genetic Technology,		
CC	Applied Biosystems Division of Perlin Elmer Corp., 950 Lincoln		
CC	Center Drive, Foster City, CA 94004 USA and David Schlessinger,		
CC	Department of Molecular Microbiology and Center for Genetics in		
CC	Medicine Washington University School of Medicine, St. Louis MO		
CC	63110 USA e-mail: ellson@genseq.aplbio.com and		
CC	david@genetics.wustl.edu Note: Gene predictions were accomplished		
CC	with runs of Grail versions 1.1 and 1.2, coupled with fasta and		
CC	blastx comparisons to genbank & non-redundant peptide libraries.		
CC	Repeat analysis was accomplished via censor. NCBI gi: 1203968		
FH	Key	Location/Qualifiers	
FT	source	1..219447	
FT		/organism="Homo sapiens"	
FT		/sequenced_mol="DNA"	
FT	repeat_unit	4..223	
FT		/evidence=EXPERIMENTAL	
FT		/map="X"	
FT	repeat_unit	/rpt_family="Alu-Sp or Alu-Sq"	
FT		226..503	
FT		/map="X"	
FT		/evidence=EXPERIMENTAL	
FT	repeat_unit	/rpt_family="Alu-J or Alu-S"	
FT		641..931	
FT		/map="X"	
FT		/evidence=EXPERIMENTAL	
FT	repeat_unit	/rpt_family="Alu-J or Alu-S"	
FT		1023..1313	
FT		/map="X"	
FT		/evidence=EXPERIMENTAL	
FT	repeat_unit	/rpt_family="Alu-Sx"	
FT		complement (1446..1536)	
FT		/evidence=EXPERIMENTAL	
FT		/map="X"	
FT	repeat_region	/rpt_family="Alu-J"	
FT		3095..3426	

Aug 27 08:17

US-08-231-565A-1.rge

35

```
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="CpG Island"
FT complement(3663..3952)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sb1"
FT 4912..5046
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT 5047..5333
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sb1"
FT 5346..5628
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT 5632..5805
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sq"
FT 6363..17698
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="11Kb repeat1"
FT 6369..6633
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT complement(7121..7409)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sp"
FT complement(7358..7432)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sb1"
FT complement(7426..7701)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT 8306..8592
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT complement(8863..9029)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(9031..9317)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sq"
FT complement(9318..9442)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT 9486..9773
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT 9996..10132
```

Aug 27 08:17

US-08-231-565A-1.rge

36

```
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sp"
FT 10159..10255
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sq"
FT 10271..10367
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT complement(10455..10739)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT complement(10740..10860)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(10866..11156)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT 11223..11501
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT 11638..13564
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="CpG Island"
FT complement(11884..12163)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sb1"
FT complement(12195..12482)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(12668..12945)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(18360..18648)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0"
FT complement(18984..19488)
FT /evidence=EXPERIMENTAL
FT /gene="FLN"
FT /map="X"
FT /number=48
FT complement(join(19301..19488,19814..20017,20101..20319,
20483..20659,21361..21493,22033..22148,22336..22473,22633,
22899,23005..23127,23224..23376,23453..23656,23748..23909,
24006..24179,24367..24495,24603..24743,24833..24935,25068,
25163,25277..25524,27701..27724,27884..28073,28649..28805,
28895..29018,29434..29604,29696..29856,29934..30096,30182,
30355,30440..31037,31758..32020,32120..32237,32429..32598,
32692..32782,32868..33028,33111..33234,34472..34615,34709,
```

```

FT FT 34822.34976..35169.35271..35407.35586..35709.35798...35933.
FT FT 36473..36673.36757..36919.37011..37088.37181..37299.37846.
.
FT FT 37993.38090..38187.38291..38539.41319..41691))
FT FT /evidence=EXPERIMENTAL
FT FT /gene="FIN"
FT FT /map="X"
FT FT /codon_start=1
FT FT /function="binds actin"
FT FT /product="filamin"
FT FT /db_xref="PID:gl203969"
FT FT /translation="MSSSHRGAQSAAGAPGGGVYDTRDAEMPTEKDIAEDAPKKI

```

...
Note: remainder of annotations omitted.

Query Match	14.9%;	Score 232;	DB 10;	Length 219447;
Best Local Similarity	90.7%;	Pred. No. 1.19e-100;		
Matches	264;	Conservative 0;	Mismatches 26;	Indels 1; Gaps 1;
Db	65611	agccgggcggtggctcacactgttaattccagcaccattggaggtcgaggcggcgaga	65670	
Cp	1259	AGCCAGGCGTGGTGGCTCACGCCCTGTAAATCCAGCATTGGGAGCGCTGAGGCGGGCAGA	1200	
Db	65671	tcacctgaggtcgaggtttgagaccagcgtgacca-cacggagaaacccctgtctctact	65729	
Cp	1199	TACCTGAGGTCGAGGTTTGAGACACGCCGTGACCAACATGAGAAACCCGCTCTCTACT	1140	
Db	65730	aaaaatcaaaaataagccaggcatggtggcacatgctctgatccagctactcgggagg	65789	
Cp	1139	AAAACTCAAAAATTAGTCAGGCATAGTGGCGCAGCCTGTAAATCCAGCTACTCTCAGAGG	1080	
Db	65790	ctgaggcaggagaattgcttgaaccgggagggcagaggtttgcaactgagctgagatcgtgc	65849	
Cp	1079	CTAAGCGAGGAGAAATTGCTTGAACTGGGAGCGGAGGTTATGCTGAGCCAAATCGCGC	1020	
Db	65850	catggcactccaagctgggcaacaagagcgaaactccatctcagaaaaaaa	65900	
Cp	1019	CATTGCACTCCAGCTGGGCAACAAAGCGAAACTCCATCTCAAAAAAAA	969	

```

RESULT 14
LOCUS          1171 bp      DNA
DEFINITION    Human Pro-TRNA and Val-TRNA genes.
ACCESSION     M27835
NID           g339615
KEYWORDS      transfer RNA-Pro; transfer RNA-Val.
SEGMENT       2 of 2
SOURCE        Human HeLa cell DNA, clone pHPTV5.
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 1171)
AUTHORS      Shortridge,R.D., Johnson,G.D., Craig,I.L.C., Pirtle,I.L. and
              Pirtle,R.M.
TITLE         A human TRNA gene heterocluuster encoding threonine, proline and
              valine tRNAs
JOURNAL       Gene 79 (2), 309-324 (1989)
MEDLINE      90006780
COMMENT      NCBI gi: 339615
FEATURES     Location/Qualifiers
              1..1171
              /organism="Homo sapiens"
              241..312
              trna

```

[illegible]

RESULT	15	
AC	HS114748C standard; DNA; PRI; 1846 BP.	
ID	Z69375;	
DT	08-FEB-1996 (Rel. 46, Created)	
DT	08-FEB-1996 (Rel. 46, Last updated, Version 1)	
DE	Human DNA sequence from cosmid L174G8, Huntington's Disease Region,	
DE	chromosome 4p16.3	
KW	4p16.3.	
OS	Homo sapiens (human)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.	
RN	[1]	
RP	1-1846	
RA	Dodsworth S., Baron L.;	
RT	;	
RL	Submitted (07-FEB-1996) to the EMBL/GenBank/DBJ databases.	
RL	Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, UK. E-mail	
RL	enquires: humquery@sanger.ac.uk	
CC	IMPORTANT:	
CC	This sequence is not the entire insert of clone L174G8. It may be	
CC	shorter because we only sequence overlapping sections once, or	
CC	longer because we arrange for a small overlap between neighbouring	
CC	submissions.	
CC	This sequence has been finished according to sequence map criteria	
CC	as follows. An attempt is made to resolve all sequencing problems,	
CC	such as compressions and repeats, but not necessarily within known	
CC	annotated human repeat sequence elements (e.g. Alu). Where the	
CC	sequence is ambiguous, there is an annotation using the "unsure"	
CC	feature key.	

```

CC The sequence from clone L174G8 has been finished in more than one
CC contig.
CC This sequence (L174G8C) is separated from the preceeding one
CC (L174G8B)
CC by a gap of about 1.2kb sized by restriction analysis.
CC The true right end of clone L174G8 is at 1846.
CC L174G8 is from cosmid library LA04NC01 constructed at the
CC Human Genome Center, Los Alamos National Laboratory, NM 87545
CC under the auspices of the U.S. Department of Energy.
CC The library was constructed using flow-sorted human chromosome 4
CC from a Hamster-Human hybrid cell line (UV20HL21-27) containing
CC human chromosomes 4, 8 and 21.
CC VECTOR: sCos1
CC L174G8 is contained in a clone contig spanning ~2Mb which is
CC described in Baxendale et al, Nature Genetics 4 ( 1993 ) 181-
CC 186. See also Myers et al, Cytogenet Cell Genet. 66 ( 1994 )
CC 218-230.
FH Key Location/Qualifiers
FT source 1..1846
FT /organism="Homo sapiens"
FT /chromosome="4"
FT /map="4p16.3"
FT /clone="L174G8"
FT /clone_lib="LA04NC01"
FT /cell_line="UV20HL21-27"
FT repeat_region 969..1203
FT /note="Alu repeat: matches 60..308 of consensus"
FT /partial
FT repeat_region 1255..1546
FT /note="Alu repeat: matches 308..1 of consensus"
FT /partial
FT repeat_region 2..272
FT /note="Alu repeat: matches 287..1 of consensus"
FT /partial
SQ Sequence 1846 BP; 530 A; 440 C; 371 G; 505 T; 0 other;
Query Match 14.8%; Score 230; DB 10; Length 1846;
Best Local Similarity 89.9%; Pred. No. 1.39e-99;
Matches 259; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Db 1254 tcttttttttgagatggagtttctctgtgtgccaggctggagtgcaatggcacgat 1313
|||||
Qy 967 TCTTTTTTTTGAGATGGAGTTTGGCTTTTGTGGCCAGGCTGGCAATGGCGCGAT 1026

Db 1314 ctcggtcactgcaacctccgctcccggttcaagccattctctgcagcctctg 1373
|||||
Qy 1027 CTTGCTCACCATRACCTCCGCCCTCCAGGTTCAAGCAATTCTCTGCTTAGCTCTG 1086

Db 1374 agtgcgtggtattacagcgccacagccaccgcggctaatgttatttttagtagag 1433
|||||
Qy 1087 AGTACCTGGGATTACAGCGGTGGCGCACTATGCTCACTAATTTGTAGTTAGTAGAG 1146

Db 1434 acagggtttctccatgttgcaggctgtctcgaactcccgacctcaaggatccgct 1493
|||||
Qy 1147 ACGGGTTTCTCCATTTGCTGAGGCTGGTCTCAAACTCCTGACCTCAGGTATCTGCC 1206

Db 1494 gctcagctcccaagtgcgtggattacaggcatgagccacgcgc 1541
|||||
Qy 1207 GCTCAGCTCCCAAGTGTGGAATTACAGGCTGAGCCACACGCC 1254

```

Search completed: Tue Aug 27 08:29:27 1996

Job time : 1307 secs.

Result No.	Query			DB	ID	Description	Pred. No.
	Score	Match	Length				
c	1	1559	100.0	1559	17	MART-1 melanoma anti	0.00e+00
	2	633	40.6	760	17	Tumor rejection anti	0.00e+00
	3	350	22.5	354	12	Tumour rejection anti	2.45e-170
	4	304	19.5	13585	17	Tumour rejection anti	3.98e-145
	5	232	14.9	59275	15	Human thympoietin co	6.55e-106
	6	215	13.8	370	8	Human brain Expressed	1.04e-96
	7	215	13.8	387	8	Human brain Expressed	1.04e-96
	8	212	13.6	8174	2	GDP-Fuc:beta-D-galact	4.32e-95
	9	212	13.6	8174	9	DNA encoding a glycos	4.32e-95

RESULT	1
ID	T02714 standard; cDNA; 1559 BP.
AC	T02714;
DT	20-APR-1996 (first entry)
DE	MART-1 melanoma antigen.
KW	MART-1; melanoma antigen recognised by T-cells; melanoma; metastatic melanoma; tumour-associated antigen; immunogen; diagnosis; prognosis; prophylaxis; therapy; vaccine; ds. Mammalian.
Key	Location/Qualifiers
CDS	54..410
FT	/!tag= a
PN	W09529193-A2.
PD	02-NOV-1995.
Pf	21-APR-1995; US05063.
PR	22-APR-1994; US-231565.
PR	05-APR-1995; US-417174.
PA	(USSH) US SEC DEPT HEALTH.
PI	Kawakami Y, Rosenberg SA;
DR	WPI; 95-382963/49.
DR	p-RPDB; R84212.
PT	DNA encoding melanoma antigens recognised by T-lymphocytes - also vectors, host cells and antibodies, used to detect, treat and immunise animal against melanoma

PS Claim 2; Page 115-116; 184pp; English.

CC The nucleic acid encodes a melanoma antigen (MART-1) which is
CC recognized by T-lymphocytes. It is used for recombinant protein
CC production, preferably using a baculo virus vector for expression in
CC insect cell cultures. MART-1 protein is a source of immunogenic
CC peptides (see R84196 for peptide M9-2) which are optionally modified
CC (see R84783-R84800) and used in medicaments for the treatment or
CC prevention (by immunization) of melanoma. Antibodies against MART-1
CC and its immunogenic peptides may be used in the detection and
CC isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 1559 BP; 470 A; 330 C; 324 G; 435 T;

Query Match 100.0%; Score 1559; DB 17; Length 1559;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 accagacagagctctcatiaaagagtgctcctgtgccctgacctcacaagatgccaa 60
Qy 1 AGCAGACAGAGACTCTCATTAAGGAAGGTGCTGTGCCCTGCACCTACAAGATGCCAA 60
Db 61 gagaagatgctcactcatctatgtttaccocaaagaaggggcaagggccactctacacca 120
Qy 61 GAGAGATGCTCATTTCATCTATGTTATCCCAAGAGGGGCGCGCACTCTTACACCA 120
Db 121 cggctgaagagcgctggagcgcatcctgacagtgcctggagctctactgtctca 180
Qy 121 CGGCTGAAGAGCGGCTGGGATCGGCATCCTGACATGATCTGGGAGTCTTACTGCTCA 180
Db 181 tcggctgttggtatgtagaagacaaatggtacagagccttgatgataaaagtctc 240
Qy 181 TCGGCTGTGTGATATGTGAAGACAGAAATGGATACAGAGCCCTTGATGAATAAAGTCTTC 240
Db 241 atgttgcactcaatgtccttaacaaagatgcccaagaaggggttgatcatcagg 300
Qy 241 ATGTGGCACTCAATGTGCTTAAACAAGAGTCCCAACAAGAGGGTTGATCATCGGG 300
Db 301 acagcaagtgtctcttcaagagaacagtgaacctgtggtcccaatgtccacactg 360
Qy 301 ACAGCAAGTGTCTCTTCAAGAGAAACTGTGAACCTGTGTGCCANTGCTCCACCTG 360
Db 361 cttatgagaactctctgagaacagtcaccacacacttattcacccttaagagccagca 420
Qy 361 CTATGAGAAACTCTCTGAGACAGCTGACACACCTTATTCACCTTAAGAGCCACCA 420
Db 421 gacactgagacatgctgaatatttctcaccactttgtgttgaaatttaacagaca 480
Qy 421 GACACCTGAGACATGCTGAAATTAATTTCTCTCACACTTTGCTTGAATTAATACAGACA 480
Db 481 tctaattctcctcttgaaatggtgtagaataatgcaagccatctcctaataaagca 540
Qy 481 TCTAATGTTCTCCTTTGGAATGTTGAGAAATGCAAGCAATCCAGCCATCTCTAATAAAGTCA 540
Db 541 gtgttaaatatttagtggtccgtagcagactaatcatgtgaggaatgatgagaat 600
Qy 541 GTGTTAAATTTTAGTGTGCTCGGTAGCAGTACTTAATCATGTGAGGAATGATGAGAAAT 600
Db 601 attaaatggaaactccatcaataaattgttgcattgcattgataactatctgtgccag 660
Qy 601 ATTAATTTGGGAAACTCTCCATCAATAAATGTTGCAATGCATCATCTGTGCGCAG 660
Db 661 gtaatgttaataatccatggtgtattttctgagagacagaattcaagtgggtattctg 720
Qy 661 GTAATGTAGTAAATCCATGTTGTTATTTTCTGAGACAGAGAAATTCAGTGGGTATTCTG 720

Db 721 ggccatccaattctcttacttgaatttggctaatacaaaactagtcagggttttca 780
Qy 721 GGGCCATCCAAATTTCTTTTACTTTGAAATTTGGCTTAATAACAACTAGTCAGGTTTCA 780
Db 781 acctgacgcagatgaactgtacacagaattgttccagttactatggagtgtccacaagg 840
Qy 781 ACCTTGCCGACATGAACGTGTACACAGAAATTTGTTCCAGTACTATGGAGTGTCAAAAG 840
Db 841 atactttacaggttaagcaaaaggttgaactggcctatttattctgatcaagaacatgtc 900
Qy 841 ATACTTTACAGGTTAAGCAAAAGGTTGACTGGCTATTATCTGATCAAGACATGTC 900
Db 901 agcaatgtctcttctgtcttaaaattctattatacaataatatttgaagaatcc 960
Qy 901 AGCAATGTCTCTTTGTGCTTAAMAATTTCTATTACTACAATAATATTGTAAAGATCC 960
Db 961 tatagctcttt 1020
Qy 961 TATAGTCTCTTTTTTTTTCAGATGAGGTTTCGCTTTTGTGCCCAGGCTGGAGTGAAT 1020
Db 1021 cggatcttggctcaccataaactccgctccaggttccaggaattctcctgccttagc 1080
Qy 1021 CGGATCTTGGCTCAGCATTAACCTCGCGCTCCAGGTTCAAGCAATTTCTCTGCTTAGC 1080
Db 1081 ctectgagtagctgggattacagcggtgcgcactatgcttgaactaattttgtagtctta 1140
Qy 1081 CTCCTGAGTAGCTGGGATTAAGCGGTGGGCCACTATGCTGACTAATTTTGTAGTTTA 1140
Db 1141 gttagagcgggttttctccatgttggcaggtgtgtctcaaaactcctgacctcagggtgat 1200
Qy 1141 GTAGACGGGGTTTCTCCATGTTGGTCAAGGCTGGTCTCAAACTCTTGACCTCAGGTGAT 1200
Db 1201 ctgcccgcctcagcctcccaagtgtcggaaattacaggggtgagccaccaagcctggctg 1260
Qy 1201 CTGCGCGCTCAGCCTCCCAAGTGTGGAATTAACAGGCTGAGCCACCAAGCCTGGCTG 1260
Db 1261 gatcctatatcttagtgaagacataaacgcagtcctaatcacatttcaactcaaggctca 1320
Qy 1261 GATCCTATATCTTAGGTAAGACATAAAGCAGTCTAATACATTTTCACTTCAAGGCTCA 1320
Db 1321 atgctattcctaactaatgacaagtattttctactaaaccagaaattgttagaaggattta 1380
Qy 1321 ATGCTATTCTAATGACAAAGTATTTTCTACTTAACCCAGAAATTTGGTAGAGGATTTA 1380
Db 1381 aataagtaaaagctactatgtactgccttagtgcctgtgactgcctttaaagt 1440
Qy 1381 AATAAGTAAAAAGCTACTATGTACTTGGCTTAGTGTGATCCCTGTGTACTTAAATGT 1440
Db 1441 acctatggcaatttagctctctgttcccaaatcctctccacaagaatgtgcagaaga 1500
Qy 1441 ACCTATGGCAATTTAGCTCTCTTTGGGTTCCCAATCCCTCTCAACAAGATGTGCAAGA 1500
Db 1501 aatcataaagatcagagatctctgggttcccaaatcctctccacaagaatgtgcagaaga 1559
Qy 1501 AATCATAAAGGATCAGAGATCTCTGAATAAATAAATAAATAAATAAATAAATAA 1559

RESULT 2

ID T11550 standard; cDNA; 760 BP.
AC T11550;
DT 18-APR-1996 (first entry)
DE Tumour rejection antigen precursor coding sequence.
KW Tumour rejection antigen precursor; TRAP; TRA; melanoma; cancer;

Db 121 ctgtctcatcgctgtgtgtatttagaagacgaatggatcacagcccttgatgataaa 180
 |||
 Qy 174 CTGCTCATCGCGCTGTGGTATTGTACAGACGAATGGATACAGCGCTTGATGATAA 233
 |||
 Db 181 agtctcatgtgtgacacaaatgtccttaacagaagatgccacaagaagggtttgat 240
 |||
 Qy 234 AGTCTTCATGTGGCACTCAATGTGGCTTTAACAAGAGATGCCACAGAAGGGTTGTAT 293
 |||
 Db 241 catcgggacagcaaaatgtctcttcaagagaaaactgtgaacctgtgttccaatgct 300
 |||
 Qy 294 CATCGGACAGCAAAAGTGTCTCTTCAAGAGAAAAGTGTGAACCTGTGGTTCCTCAATGCT 353
 |||
 Db 301 ccacctgttatgagaactctctgcagaacagtcacaccacacctattcaact 354
 |||
 Qy 354 CCACCTGCTTATGAGAACTCTCTGCAGACAGCTGCACACCACTTATTCACCT 407
 |||

RESULT 4

ID T11549 standard; DNA, 13585 BP.
 AC T11549;
 DE Tumour rejection antigen precursor coding sequence.
 KW Tumour rejection antigen precursor; TRAP; TRA; melanoma; cancer;
 KW tumour; treatment; detection; vaccine; HLA-A2; adoptive transfer;
 KW T cell; T lymphocyte; human leukocyte antigen; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT misc feature 2685
 FT /*tag= a
 FT /note= "Unidentified nucleotide."
 FT misc feature 9422..9456
 FT /*tag= b
 FT /note= "This region is 4.7-5.3 kilobases in length
 FT but the sequence of this region has not been
 FT deduced."
 FT misc feature 11539
 FT /*tag= c
 FT /note= "Unidentified nucleotide."
 PN W09601557-A1.
 PD 25-JAN-1996.
 PF 27-JUN-1995; U08153.
 PR 08-JUL-1994; US-212351.
 PR 10-JAN-1995; US-370319.
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Boon-falleur T, Brichard V, Coulie P, De PLAEN E;
 PI Traversari C, Van PEL A, Wolfel T;
 DR WPI; 96-097390/10.
 PT Genomic DNA encoding a tumour rejection antigen precursor -
 PT processed to antigen presented by HLA-A2, useful for treating or
 PT diagnosing melanoma
 PS Claim 3; Page 18-22; 41pp; English.
 CC The tumour rejection antigen precursor (TRAP) is processed to a
 CC tumour rejection antigen (TRA) presented by HLA-A2 molecules. TRA
 CC is used to generate cytotoxic T lymphocytes for treating cancer (esp.
 CC melanoma). It can also be used to raise specific antibodies, and
 CC when complexed with HLA-A2, it can be used to produce vaccines.
 CC Cytotoxic T lymphocytes so generated can be used in adoptive
 CC transfer or generated, or they can be generated in vivo by using a
 CC vector containing the appropriate gene or using TRA or TRAP together
 CC with an adjuvant that facilitates entry into HLA-A2 presenting cells.
 CC Diagnostic methods involving the detection of expression of TRAP
 CC can be used in the detection of cancers.
 SQ Sequence 13585 BP; 3827 A; 2859 G; 2968 G; 3894 T;

Query Match 19.5%; Score 304; DB 17; Length 13585;
 Best Local Similarity 100.0%; Pred. No. 3.98e-145;
 Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 13274 gttcccaatgctccacgtcttatgagaaactctctgcagaacagtcacaccaccttat 13333
 |||
 Qy 342 GTTCCCAATGCTCCACCTGCTTATGAGAAACTCTCTGCAGACAGTCACACCACCTTAT 401
 |||
 Db 13334 tcaccttaagagccagcagacacacctgagacatgctgaaattattctctcacactttg 13393
 |||
 Qy 402 TCACCTTAAGAGCCAGCCAGACACCTGACACATGCTGAAATTTATTTCTCACACTTTG 461
 |||
 Db 13394 cttgaatttaatacagacaactaatgtctctctttggaatggtgtaggaataatgcaagc 13453
 |||
 Qy 462 CTTGAATTTAATACAGACATCTAATGTTCTCTTTGCAATGCTGTAGGAAAATGCAAGC 521
 |||
 Db 13454 catctctaataaagtcagtggttaaaatttagtgctcgctagctactaatcatg 13513
 |||
 Qy 522 CATCTCTAATAAAGTCACTGTTAAATTTTAGTGGCTAGCAGTACTAATCATG 581
 |||
 Db 13514 tgaggaatgatgagaatattaaattgggaaaaactccatcaataaattgtgcaatgcat 13573
 |||
 Qy 582 TGAGGAATGATGAGAAATATTAAATTTGGGAAAACCTCCATCAATAATGTTGCAATGCAT 641
 |||
 Db 13574 gata 13577
 |||
 Qy 642 GATA 645
 |||

RESULT 5

ID Q92779 standard; DNA, 6905 BP.
 AC Q92779;
 DT 09-DEC-1995 (first entry)
 DE Human thymopoietin continuous gene fragment.
 KW Thymopoietin; diagnostic; therapeutic; gene therapy; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 39..11956-2843
 FT /*tag= a
 FT /note= "exon 4"
 FT exon 4691..4788
 FT /*tag= b
 FT /note= "exon 5"
 PN W09517205-A1.
 PD 29-JUN-1995.
 PF 13-DEC-1994; U14356.
 PR 21-DEC-1993; US-171382.
 PA (IMMU-) IMMUNOBIOLOGY RES INST INC.
 PI Culler MD, Goldstein G, Harris CA, Setcavage DR;
 PI Shenbagamurthi P, Siekierka JJ, Talle MA;
 DR WPI; 95-240474/31.
 PT Polynucleotide(s) encoding human thymopoietin proteins - used to
 PT develop prods. for diagnosis and therapy involving immune or nervous
 PT system conditions.
 PS Disclosure; Fig 9a-9g; 85pp; English.
 CC The sequence represents a fragment of the human thymopoietin gene,
 CC including exons 4 and 5. DNA encoding human thymopoietin, from
 CC a cDNA library prepared from human thymus RNA, may be expressed in
 CC a host cell e.g. bacterium (preferably Escherichia coli), fungus,
 CC insect or mammalian cell for production of recombinant thymopoietin.
 CC Thymopoietin has a regulatory effect on the mammalian immune system,
 CC and can be used for treating/modulating an immune or nervous system
 CC condition, and for the treatment of chronic infection, autoimmune
 CC disorders and certain affective psychiatric or neurological

QY 1106 GTGCCCCACTATGCTCGTCAATTTTGTAGTTTGTAGTACAGACGGGGTTTCTCATGTTG 1165
 Db 180 gtcaggctggtcttgaactcctgaactcagtgatccgctgctcgcgctcccaagtgc 239
 QY 1166 GTACGGCTGGTCTCAAACTCTGACCTCAGGTGATCTGCCGGGCTCAGCTCCCAAGTG 1225
 Db 240 ctggaggattacaggcataagccactgagcc 268
 QY 1226 CTGGAATTACAGGCTGTAGCCCAAGCGC 1254

RESULTS

ID	Q13332	standard; DNA; 8174 BP.
AC	Q13332;	
DT	07-NOV-1991	(first entry)
DE	GPB-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase gene.	
KW	Glycosyltransferase.	
OS	Homo sapiens.	
EH	Key	Location/Qualifiers
FT	CDS	4686..5783
FT	FT	/*taq= a
PN	W09112340-A.	
PD	22-AUG-1991.	
PR	14-FEB-1991;	U00899.
PR	14-FEB-1990;	US-480133.
PR	14-FEB-1990;	US-479858.
PR	12-DEC-1990;	US-627621.
PA	(UNMI) UNIV OF MICHIGAN.	
PI	Lowe JB;	
DR	WPI; 91-267151/36.	
DR	P-PSDB; R13751.	
PT	Isolation of gene conveying post-translational characteristic -	
PT	e.g. the presence of soluble or membrane bound oligo or	
PT	polyaccharide or glycosyltransferase.	
PT	Disclosure; Fig 3; 155pp; English.	

The DNA encodes a protein sequence capable of functioning as a GDP- Fuc: [beta-D-Gal alpha(1,2)- fucosyltransferase. The sequence coded by nucleotides 4782 - 5780 represents the functional protein. The enzyme produced by the DNA sequence can be used in enzymatic fucosylation of chain-terminating galactose residues on lactose-amine or neolacto type beta-D-galactoside to alpha-2-L-fucose residues. See also Q13330-Q13333.

Query Match	13.6%	Score 212;	DB 2;	Length 8174;
Best Local Similarity	86.6%	Pred. No. 4.32e-95;		
Matches	251;	Conservative	0;	Mismatches 39; Indels 0; Gaps

[illegible]

Db 4067 cagctcaccgcaactccgcgcgcgcgggttcaggcgattctctgcctcagctcccgag 4126

Qy 1029 TGGCTCACCATAACCTCCGCTCCAGGTTCAAGCAATCTCTGCGCTTAGCCTCCTGAG 1088

Db 4127 TGGCTGGGATTACAGGCGATCGCCACCATGCGCGGCTAATTGTATTTAAGTAGAGAC 4186

Qy 1089 TAGCTGGGATTACAGGCGTGGCCCACTATGCCCTGACTAATTTTGTAGTTTCTAGAGAC 1148

[illegible]

Db

4247 ctggcctcccaagtgcgtgattacaggtgatgccaccgcgcctggc 4296
||| ||||||||
Gv

1209 CTCAGCCTCCCAAGAATGCTGGATTACAGCGGTGAGGCCAACCCTTGC 1258

RESULTS

ID	Q56908	standard; DNA; 8174 BP.
AC	Q56908;	
DT	26-JUL-1994	(first entry)
DE	DNA encoding a glycosyltransferase.	
KE	Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;	
KW	surface; oligosaccharide; ss.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
CDS	FT	4886..5783
PN	/*tag= a	
PD	W09402616-A.	
PP	03-FEB-1994.	
PR	20-JUL-1993; U06703.	
PA	20-JUL-1992; US-914281.	
PI	(UNMI) UNIV MICHIGAN.	
DR	Lowe JB;	
DR	WPI; 94-048874/06.	
DR	P-PSDB; R45936.	
PT	DNA fragment encoding a glycosyltransferase - can be used for in vitro reactions to modify cell surface oligosaccharide(s) e.g. blood gp. determinants, to protect against transplant rejection	
PT	Disclosure; fig 3; 249pp; English.	
CC	The sequence is that encoding human glycosyl transferase. The enzyme produced by the DNA may be non glycosylated. This prevents premature loss of enzyme activity. It can also be used in in vitro reactions to modify cell surface oligosaccharide mols. e.g. blood group determinants.	
CC	See also Q56905-12.	
CC	Sequence	8174 BP; 1628 A; 2228 C; 2322 G; 1996 T;

Query Match 13.6%; Score 212; DB 9; Length 8174;
Best Local Similarity 86.6%; Pred. No. 4.32e-95;
Matches 251; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

[illegible]

Db 4067 cagctcaacgcaactccgcgcgcgggttcaggcgattctctgcctcagctcccgag 4126

Qy 1029 TGGGTACACCA⁺AACCTCCGCTCCAGGTTACAGCAATTCTCCTGCCTTAGCCTCCTGAG 1088

Db 4127 tggctggattacaggcatggccaccatgccgggctaattttgtatttttaagt⁺agagac 4186

Qy 1089 TAGCTGGGATTACAGSGGTGGCCACTATGCCTCACTAATTTTGTAGTTTGTAGTAGAGAC 1148

DB 418/ agggttcttcacgttggtcaggtggtctgaactcccaacctcaggtgatccaccac 4248
|||||
Qy 1149 GGGGTTTCTCCATGTTGGTCAGGCTGGTCTCAAACTCCTGACCTCAGTGATCTGCCCG 1208

Db 4247 cttggcctcccaaatgctgggattacaggtgtgagccaccgagcctggc 4296
 || |||||
 Qv 1209 CTCAGCCTCCCAAGTCTGTGAAATACAGCGTGAGCACCACGCCTGGC 1258

RESULT 10

ID	N81564 standard; DNA; 11715 BP.
AC	N81564;
DT	05-DEC-1990 (first entry)

DE Genomic sequence encoding protein C.
KW Human protein C; blood coagulation disorders; ss.
FH Key Location/Qualifiers
FT exon 2131..2200
FT /*tag= a 3464..3630
FT /label=exon I
FT exon
FT /*tag= b 5093..5117
FT /label=exon II
FT exon
FT /*tag= c 5212..5349
FT /label=exon III
FT exon
FT /*tag= d 5452..5586
FT /label=exon IV
FT exon
FT /*tag= e 8255..8397
FT /label=exon VI
FT exon
FT /*tag= f 9270..9387
FT /label=exon VII
FT exon
FT /*tag= g 10517..12006
FT /label=exon VIII
FT /*tag= h
FT /label=exon IX
PN EP-266190-A.
PD 04-MAY-1988.
PR 28-OCT-1987; 309528.
PR 29-OCT-1986; US-924462.
PA (Zymo-) Zymogenetics Inc.
PI Foster DC, Murray MJ, Berkner KL.;
DR WPI; 88-121259/18.
DR P-PSDB; P81205.
PT Protein C DNA coding sequence and expression vector for prodn. -
PT used for treating blood coagulation disorders.
PS Disclosure; p; English.
CC In the construction of the full length protein C gene this sequence
CC the introns removed then is joined to a cDNA clone. The desired
CC genomic:cDNA is then generated by looping out of unwanted sequences
CC using oligonucleotide-directed deletion mutagenesis. The protein
CC produced upon transformation of mammalian host cells, contg. the
CC recombinant DNA, has substantially the same biological activity
CC as natural protein C and is hence useful in the treatment of blood
CC coagulation disorders. See also N81563.
SQ Sequence 11715 BP; 2443 A; 3292 G; 3375 G; 2605 T;
Query Match 13.6%; Score 212; DB 1; Length 11715;
Best Local Similarity 87.5%; Pred. No. 4.32e-95;
Matches 253; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
Db 6376 tcttttttttgatgagtttcaactcttctgtccagctgagtgcaatgacgtga 6435
|||||
QY 967 TCTTTTTTTTGTAGATGGAGTTTCGCTTTCTTG-CCACAGCTGGAGTGCATGGCGCA 1025
Db 6436 tctcagctcaccacaacctccgctcctggattcaagcgattctcctcgcgagctccc 6495
|||||
QY 1026 TCTTGCTCACCATAACCTCGCGCTCCAGGTTCAAGCAATTCCTCGCTTAGCCTCT 1085
Db 6496 gadtgctgggattacagagcatgcgaccacgcagctaatttgtgttttagtaga 6555
|||||
QY 1086 GAGTAGCTGGGATTACAGCGCTGCGCCACTATGCTGACTAATTTCTAGTTTAGTAGA 1145

Db 6556 gaagggtttctcgtgtgttgcgaagctgttgaactcctgacctcaggtgatccacc 6615
|||||
QY 1146 GACGGGGTTTCTCATGTGCTGAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCTGCC 1205
|||||
Db 6616 tgcttggcctcctaaagtctggtgattacagggctgagccaccgcgc 6664
|||||
QY 1206 CGGCTCAGCTCCCAAGTGTCTGAATTACAGGGCTGAGCCACCAAGCC 1254
|||||
RESULT 11
ID N70102 standard; DNA; 11724 BP.
AC N70102;
DT 10-MAY-1991 (first entry)
DE Complete genomic sequence of human Protein C.
KW Human Protein C; anti-coagulant; thrombosis; serine protease; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT exon 2131..2200
FT /*tag= a
FT /number= 1
FT /note= "encodes amino acids -42 to -20"
FT intron 2201..3463
FT /*tag= b
FT /number= 1
FT exon 2464..3630
FT /*tag= c
FT /number= 2
FT /note= "encodes amino acids -19 to +37"
FT intron 3631..5092
FT /*tag= d
FT /number= 2
FT exon 5093..5117
FT /*tag= e
FT /number= 3
FT /note= "encodes amino acids 38 to 45"
FT intron 5118..5209
FT /*tag= f
FT /number= 3
FT exon 5210..5347
FT /*tag= g
FT /number= 4
FT /note= "encodes amino acids 46 to 91"
FT intron 5348..5449
FT /*tag= h
FT /number= 4
FT exon 5450..5584
FT /*tag= i
FT /number= 5
FT /note= "encodes amino acids 92 to 136"
FT intron 5585..8252
FT /*tag= j
FT /number= 5
FT exon 8253..8395
FT /*tag= k
FT /number= 6
FT /note= "encodes amino acids 137 to 184"
FT intron 8396..9267
FT /*tag= l
FT /number= 6
FT exon 9268..9385
FT /*tag= m
FT /number= 7
FT /note= "encodes amino acids 185 to 223"
FT intron 9386..10514

FT /*tag= n
 FT /number= 7
 FT exon
 FT /*tag= o
 FT /number= 8
 FT /note= "encodes amino acids 224 to 419"
 FT polyA signal 11151..11156
 FT /*tag= q
 FT polyA site 11173
 FT /*tag= i
 FT polyA signal 11380..11385
 FT /*tag= g
 FT polyA site 11399
 FT /*tag= t
 PN EP-215548-A.
 PD 25-MAR-1987.
 PF 26-JUN-1986; 304970.
 PR 27-JUN-1985; US-749600.
 PR 15-AUG-1985; US-766109.
 PA (ZYMO-) ZYMOGENETICS INC.
 PA (UNIW) UNIV OF WASHINGTON.
 PI Murray MJ, Berkner KL, Foster DC, Davie EW.
 DR WPI; 87-081505/12.
 DR P-PSDB; P70855.
 PT Human protein C or activated protein C - prep. using expression
 PT vector capable of integration in mammalian host cell DNA
 PS Claim 2; Page 33; 52pp; English.
 CC A lambda gt cDNA library was prepared from human liver mRNA. The
 CC library was screened with iodine-125 labelled antibody to human
 CC protein C. A cDNA which lacked the coding sequence for the pre-pro
 CC peptide and the first 23 amino acids of protein C was isolated and
 CC used as a hybridisation probe. The remainder of the coding sequence
 CC was obtained from a human genomic DNA library in lambda Charon 4A.
 CC Recombinantly produced protein C can be used to treat thrombotic
 CC disorders such as venous thrombosis.
 SQ Sequence 11724 BP; 2445 A; 3302 C; 3370 G; 2607 T;
 Query Match 13.6%; Score 212; DB 2; Length 11724;
 Best Local Similarity 87.5%; Pred. No. 4.32e-95;
 Matches 253; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
 Db 6374 tcttttttttgagatggagtttcaactctgtgtccaggctggagtgcaatgacgtga 6433
 QY 967 TCTTTTTTTTGGATGGAGTTTCGCTTTTGTG-CCAGGCTGGAGTGCATGGCGGA 1025
 Db 6434 tctcagctcaccacaacctcgcctcctggattcaagcattctcctgcgcagctccc 6493
 QY 1026 TCTTGGCTCACCATAACCTCGGCTCCAGGTTCAAGCAATTCCTCGCTTAGCCTCT 1085
 Db 6494 gagtagctgggattcaagcagcagccaccagcagcagtaattgtgttttagtaga 6553
 QY 1086 GACTAGCTGGGATTACAGCGTGGCCACTATGCTGACTAAATTTTGTAGTTTAGTGA 1145
 Db 6554 gaaggggttctcgtgtgtgtgcaagctggcttgaactcctgacctcaggtgatcaac 6613
 QY 1146 GACGGGGTTTCCATGTTGGTCAGCGTGGTCTCAAACTCTGACCTGAGTGATCGCC 1205
 Db 6614 tgcctggcctcctaaagtctgggattacaggcgtgagccaccgcgc 6662
 QY 1206 CGCTCAGCTCCCAAGTGTGTAATTACAGGGCTGAGCCACACGCC 1254
 RESULT 12
 ID Q85367 standard; cDNA; 4788 BP.

AC Q85367;
 DT 29-AUG-1995 (first entry)
 DE MIP-1-alpha.
 KW Macrophage inflammatory protein 1-alpha; MIP-1-alpha;
 KW heparanase; heparin; heparan sulfate; arthritis; restenosis;
 KW cancer; wound healing; ss.
 OS Homo sapiens.
 PN W09504158-A.
 PD 09-FEB-1995.
 PF 26-JUL-1994; U08207.
 PR 29-JUL-1993; US-099866.
 PR 13-OCT-1993; US-136117.
 PA (UPJO) UPJOHN CO.
 PI Hoogwerf AJ, Ledbetter SR;
 DR WPI; 95-082239/11.
 DR P-PSDB; R70791.
 PT Screening for cpds. with anti-heparanase activity - by detecting
 PT inhibition of heparin or heparan sulphate degradation, cancer.
 PT potentially useful for treating arthritis, restenosis, cancer.
 PS Disclosure; Page 44-47; 60pp; English.
 CC Purified heparanases, prepared under reducing conditions and
 CC activated with transglutaminase, are given in R70786-805. Most
 CC are prepared by reverse transcription of mRNA from activated human
 CC leukocytes, then cloning of the cDNA into pVL1392 baculovirus
 CC vector, and expression in Sf9 cells in the presence of reduced
 CC glutathione and dithiothreitol.
 SQ Sequence 4788 BP; 1299 A; 1122 C; 1049 G; 1318 T;
 Query Match 13.3%; Score 207; DB 14; Length 4788;
 Best Local Similarity 87.5%; Pred. No. 2.16e-92;
 Matches 259; Conservative 0; Mismatches 34; Indels 3; Gaps 3;
 Db 1533 ttttttttgagatggagtttcaactctgtgtccaggctggagtgcaatgcatgatct 1592
 QY 969 TTTTTTTTGGATGGAGTTTCGCTTTTGTGCCCAGGCTGAGTCAAGCAATTCCTCGCTT-AGCCTCCTGA 1028
 Db 1593 cagctcaccacagcccccactcctcctgggttcaagcattctcctgcctcagctccaga 1652
 QY 1029 TGCTTCAACCAATACCTCGGCTCCAGGTTCAAGCAATTCCTCGCTT-AGCCTCCTGA 1087
 Db 1653 gtagctgggattacaggcatgtgccaccacgctgactacttttttagtagtagaga 1712
 QY 1088 GTAGCTGGGATTACAGGCTGGCCACTATGCTGACTAAATTTTGTAGTTTAGTAGAGA 1147
 Db 1713 tggagtttctcttctgtgtcaggttggctcaaacctcctgacctcaggtgatcga-g 1771
 QY 1148 CGGGGTTTTCGATGTTGGTCAGGCTGGTGTCTCAAACTCTGACCTCAGGTGATCTGCCG 1207
 Db 1772 cctcggcctcc-aaagtgtgggattacaggtgtgagcagaccatgctggctgcat 1826
 QY 1208 CCTCAGCCTCCCAAGTGTGTAATTACAGGCTGAGCCACACCGCTGGCTGCAT 1263
 RESULT 13
 ID Q54222 standard; DNA; 11531 BP.
 AC Q54222;
 DT 27-JUN-1994 (first entry)
 DE BSSL/CEL Gene.
 KW BSL; biologically functional bile salt stimulated lipase; CEL;
 KW carboxyl ester lipase; hybridisation; milk protein; transgenic;
 KW infant milk substitute; oral; specific; lipolysis; vitamin; bile;
 KW lipid malabsorption; cystic fibrosis; chronic pancreatitis;
 KW digestion; ss.
 OS Homo sapiens.

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	1653...11397	
FT	/tag= a		
FT	/product= BSSL/CEL protein		
FT	5'UTR	1..1640	
FT	/tag= b		
FT	TATA_signal	1611..1617	
FT	/tag= c		
FT	exon	1641..1727	
FT	/tag= d		
FT	exon	4071..4221	
FT	/tag= e		
FT	exon	4307..4429	
FT	/tag= f		
FT	exon	4707..4904	
FT	/tag= g		
FT	exon	6193..6323	
FT	/tag= h		
FT	exon	6501..6608	
FT	/tag= i		
FT	exon	6751..6868	
FT	/tag= j		
FT	exon	8335..8521	
FT	/tag= k		
FT	exon	8719..8922	
FT	/tag= l		
FT	exon	10124..10321	
FT	/tag= m		
FT	exon	10650..11490	
FT	/tag= n		
FT	3'UTR	11491..11531	
FT	/tag= o		
PN	W09325669-A.		
PD	23-DEC-1993.		
PF	09-JUN-1993; SE0515.		
PR	11-JUN-1992; SE-001809.		
PR	12-JUN-1992; SE-001826.		
PR	03-JUL-1992; SE-002088.		
PR	19-MAR-1993; SE-000902.		
PA	(ASTR) ASTRA AB.		
PI	Bjursell KG, Carlsson PNI, Enerback CSM, Hansson SL;		
PI	Lidberg UFF, Nilsson JA, Tornell JBF, Toernell JBF;		
PI	Enerbaeck CSM;		
DR	WPI; 94-007527/01.		
DR	P-PSDB; R45189.		
PT	DNA encoding bile salt-stimulated lipase/carboxyl ester lipase -		
PT	expressed by transgenic non-human mammals and used as human milk		
PT	substitute, contains intron sequences		
PS	Claim 1; Page 43-51; 76pp; English.		
CC	The primers (Q54223-28) are used to amplify fragments of the		
CC	BSSL / CEL gene, (Q54222). Which encodes a bile salt-stimulated		
CC	lipase / carboxyl ester lipase which when expressed by transgenic		
CC	animals can be used as a human milk substitute. The BSSL is designed		
CC	to pass through the stomach and is activated in the small intestine.		
CC	As it has a broad substrate specificity it can mediate the complete		
CC	digestion of most dietary lipids.		
SQ	Sequence 11531 BP; 2464 A; 3527 C; 3280 G; 2260 T;		
Query Match	13.2%; Score 206; DB 9; Length 11531;		
Best Local Similarity	86.0%; Pred. No. 7.47e-92;		
Matches	246; Conservative 0; Mismatches 40; Indels 0; Gaps 0;		
Db	5237 ggcgcgcgcctcacctctgtaatccaccacgcttgaggagcccaagtgaggtagatcacc 5296		

Cp	1254	GGCGTGTGGCTCAGCCCTCTAAATTTCCAGCAGCTTTTGGGAGGCGCTGAGGGGGCAGATCACC	11191
Db	5297	tgaggtcaggaagttagacacagcctgactaaaaatggaaaaacctctctctactaaaaa	5356
Cp	1194	TGAGGTGAGAGTTTGACACACGCTGACCAACTGGGAGAAACCCGCTCTCTACTTAAC	1135
Db	5357	tacaaaattagccagcgctggtgcgcgttgctgtaatccacgactactcggaggtcgag	5416
Cp	1134	TACAAAATTAGTCAGGCAATAGTGGGCGACGCTGTAAATCCGCTACTCTCAGGAGGCTAAG	1075
Db	5417	gcaggagaatcgtctgaactcagagggcggagagttgcggtgagccgagatcatgccactg	5476
Cp	1074	GCAGCAGAAATGCTTCAACTGGGAGGGCGAGGTTATGGTGAGCCAAAGTCGGCGCCATTG	1015
Db	5477	cactccagcctggagacaacagatgaaaactctgctcaaaaaaaa	5522
Cp	1014	CACCTCCAGCTGGGCAACAAAGCGAAACTCCATCTCAAAAAAAA	969

RESULT 14

ID	Q63862 standard; cDNA; 283 BP.	
AC	Q63862;	
DT	29-JAN-1995 (first entry)	
DE	AP2 sequence obtd. by PCR for tumour specific DNA.	
KW	Arbitrary primers; AP-PCR; amplification; tumour cells; cancer;	
KW	insertions; deletions; ss.	
OS	Synthetic.	
PN	W09411531-A.	
PD	26-MAY-1994.	
PF	12-NOV-1993; U10904.	
PR	13-NOV-1992; US-975737.	
PA	(CALB-) CALIFORNIA INST BIOLOGICAL RES.	
PI	Ionov Y, Malkhosyan S, McCelliand M, Peinado MA;	
PI	Perucho M, Welshj;	
WP	WPI; 94-183529/22.	
PT	Identification of tumour cells - by analysing DNA to determine	
PT	whether insertions or deletions have occurred in reiterated	
PT	sequences	
PS	Disclosure; Page 52; 67pp; English.	
CC	The sequence was obtd. by PCR with arbitrary PCR primers used to	
CC	detect insertions or deletions in DNA sequences. Such mutations are	
CC	markers of cancer so such primers can be used in the diagnosis of	
CC	cancer, esp. colorectal, stomach or pancreatic tumours.	
CC	See also Q63837-63.	
SQ	Sequence 283 BP; 63 A; 77 C; 94 G; 49 T;	

Query Match	13.1%;	Score 205;	DB 11;	Length 283;
Best Local Similarity	88.5%;	Pred. No. 2.59e-91;		
Matches 247;	Conservative	0;	Mismatches 30;	Indels 2;
Gaps 2;				
Db	6	ggcgtggtggctcacacctgaatcccagcacctttggagcgcgaggtgggtggatcacc	65	
Cp	1254	ggcgtggtggcttgcgcctgtaattccacacatttggcggcgtgagcggcgagatcacc	1195	
Db	66	tgaggtcaggagttcaagcacagcctggccaaacatggtgaaccccgctctctactaaaa	125	
Cp	1194	tgaggctcaggagtttgagcagcagcctgcaccaacatgcagaaaaccccgctctactaaaac	1135	
Db	126	tacaaaaattagccggcggtggtggcgcgccctgtaatcccagctactcgggaggtga	185	
Cp	1134	tacaaaa- tt agtcaggcctatagtcgcacgcctgttaattccacgctactcaggagcctaa	1076	
Db	186	ggcaggagaatcgcttgaacccgggaggtggagttgcagtgagccgagatcgccaccat	245	

Aug 27 08:19

US-08-231-565A-1.mg

19

Cp 1075 GCAGGAGATTCTTGAACCTGGAGGCGGAGGTTATGCTGAGCCAGATGCGGCATT 1016

Db 246 geactccagctggcgacaga-gcgagactcagctca 283

||||| ||||| ||||| ||||| ||||| ||||| |||||

Cp 1015 GCATCCAGCTGGCAACAAAGCGMAACTCCATCTCA 977

RESULT 15

ID Q59876 standard; cDNA; 257 BP.

AC Q59876;

DT 16-MAR-1994 (first entry)

DE Human brain Expressed Sequence Tag EST01565.

KW Gene transcription product; genetic markers; tagging; in vivo;

KW transcription; mapping; location; chromosomes; chromosomal; ss.

OS Homo sapiens.

PN W09316178-A.

PD 19-AUG-1993.

PF 12-FEB-1993; U01294.

PR 12-FEB-1992; US-837195.

PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

PI Adams MD, Moreno RF, Venter CJ.

DR WPI: 93-272882/34.

PT Enriched oligonucleotides and corresp. sequences - used as

PT markers for human genes transcribed in-vivo, facilitate tagging

PT of most human genes

PS Example 4; Page 246; 500pp; English.

CC The Expressed Sequence Tag was isolated from a human brain cDNA

CC library as part of a large set of ESTs which can be used as markers

CC for human genes transcribed in vivo. They can be used to facilitate

CC tagging of most human genes, for mapping locations of expressed genes

CC on chromosomes, for individual or forensic identification, for mapping

CC locations of disease-associated genes, for identification of tissue

CC type, and for prepn. of antisense sequences, probes and constructs.

CC EST01565 has a "poor" coding probability as evaluated using the

CC coding-region prediction program CRM. See also Q59041-Q61440.

SQ Sequence 257 BP; 48 A; 75 C; 60 G; 70 T;

Query Match 13.0%; Score 203; DB 8; Length 257;

Best Local Similarity 89.5%; Pred. No. 3.10e-90;

Matches 231; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

Db 1 ttccctctcgtgcccaggtgagtgcaatgctgctcancctcccagtagctgggattacaggca 120

||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 987 TTTGCTTTTGTGGCCAGGCTGGAGTGGCGGATCTTGGCTCACCATAACCTCC 1046

Db 61 gctcccaggttcaagcaattctcctgctcancctcccagtagctgggattacaggca 120

||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1047 GCCTCCAGGTTTCAGCAATTCCTGCTTACGCTCTGAGTAGCTGGGATTACAGGG 1106

Db 121 tgtncaccacgctggcctaattnta-tttaagtagagaggggtttctccatgtgg 179

||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1107 TGGCCACTATGCTGACTAATTTGTAGTTTAGTAGAGAGGGGGTTCTCCATGTGG 1166

Db 180 tcagtcctggctcaaacctcagctcaggtgactcggccacctcgccctccaaagtc 239

||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1167 TCAGGCTGGTCTCAAACTCCCTCAGCTCAGGTGATCTGCCGCCCTCAGCGCTCCCAAGTGC 1226

Db 240 tgggattacaggtgtgag 257

||||| ||||| |||||

Qy 1227 TGAATTACAGGCGTGCAG 1244

Search completed: Tue Aug 27 08:32:11 1996

Job time : 147 secs.

Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

(TM)

MPsrch_tpn n.a. - n.a. Smith-Waterman search, using a protein query
which has been backtranslated into n.a. using IUPAC symbols

Run on: Tue Aug 27 18:37:02 1996; MasPar time 260.36 Seconds
Tabular output not generated.

Title: >US-08-231-565A-2
Description: (1-118) from US08231565A.pep
Perfect Score: 1420
N.A. Sequence: 1 ATCCGNNNGARGAYGNCAC.....SNCNCNCNCTAWSNCCN 354
Comp: TACGCKNCKCTCTCCGCT.....MNGCGNGGNGATNSWNGN

Scoring table: TABLE bktranslate2
Gap 30

Nmatch STD : Dbase 0; Query 0

Searched: 270440 seqs, 380027776 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-new3
1: BCT 2: FUN 3: INV1 4: INV2 5: INV3 6: MAM 7: ORG 8: PLN
9: PRI1 10: PRI2 11: PRO 12: ROD 13: SYN 14: UNC 15: VRT
16: VIR1 17: VIR2

Database: genbank92
18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7
25: BCT8 26: BCT9 27: INV1 28: INV2 29: INV3 30: INV4 31: INV5
32: INV6 33: MAM1 34: MAM2 35: PAT1 36: PAT2 37: PAT3 38: PHG
39: PLN1 40: PLN2 41: PLN3 42: PLN4 43: PLN5 44: PLN6 45: PLN7
46: PLN8 47: PRI1 48: PRI2 49: PRI3 50: PRI4 51: PRI5 52: PRI6
53: PRI7 54: PRI8 55: PRI9 56: PRI10 57: PRI11 58: PRI12
59: PRI13 60: ROD1 61: ROD2 62: ROD3 63: ROD4 64: ROD5 65: ROD6
66: ROD7 67: STR 68: SYN 69: UNA 70: VRL1 71: VRL2 72: VRL3
73: VRL4 74: VRL5 75: VRL6 76: VRL7 77: VRT1 78: VRT2 79: VRT3

Database: genbank-new1
80: BCT 81: INV1 82: INV2 83: MAM 84: PHG 85: PLN 86: PRI
87: ROD 88: STR 89: SYN 90: UNA 91: VRL 92: VRT
93: part1

Database: u-embl45 92
93: part1

Statistics: Mean 63.591; Variance 167.878; scale 0.379

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	Pred. No.	
		Match	Length					
1	1392	98.0	752	52	HSU06654	Human differentiation	1.18e-183	
2	1392	98.0	1524	52	HSU06452	Human melanoma antigen	1.18e-183	
3	172	12.1	12857	59	S77127	manganese superoxide	9.32e-04	
4	147	10.4	41591	27	ACU12386	Acanthamoeba castella	3.52e-01	
5	142	10.0	6004	52	HSS002G	H.sapiens SOD-2 gene	1.08e+00	
6	138	9.7	1467	36	A21548	probe for the identifi	2.58e+00	
7	138	9.7	1467	22	LEPREPEL	Leptospira hardjo rep	2.58e+00	
8	137	9.6	267	50	HSG1NRS09	H.sapiens gene for gl	3.21e+00	
c	9	137	9.6	1384	28	CTBR21	Chironomus tentans Ba	3.21e+00
c	10	137	9.6	3347	77	CHKTGFBA	Gallus gallus transfo	3.21e+00
c	11	137	9.6	7471	77	BRNOTCH	B.rexio Notch mRNA.	3.21e+00
c	12	137	9.6	24537	28	CELF47F2	Caenorhabditis elegans	3.21e+00
c	13	133	9.4	1458	22	LBISLIKE	L.borgpetersenii IS-1	7.52e+00
c	14	133	9.4	3287	22	HEUNIFABX	Herbaspirillum serope	7.52e+00
c	15	132	9.3	708	34	OATCVG6	O.aries rearranged T-	9.28e+00
c	16	132	9.3	767	34	OATCCG4	O.aries rearranged T-	9.28e+00
c	17	132	9.3	2208	34	RABRAT	Oryctolagus cuniculus	9.28e+00
c	18	132	9.3	2254	34	OCRRATPRO	O.cuniculus mRNA for	9.28e+00
c	19	132	9.3	3556	71	GTVVEF	Trichoplusia ni granu	9.28e+00
c	20	132	9.3	3572	16	BAUGVEP	Pseudalatia unipuncta	9.28e+00
c	21	132	9.3	3572	91	UGVEP	Pseudalatia unipuncta	9.28e+00
c	22	132	9.3	3572	76	UGVEP	Pseudalatia unipuncta	9.28e+00
c	23	132	9.3	3934	34	RAMENGLY	Oryctolagus cuniculus	9.28e+00
c	24	132	9.3	43449	27	CELC18A11	Caenorhabditis elegans	9.28e+00
c	25	130	9.2	1599	8	BNPGALACR	B.napus mRNA for poly	1.41e+01
c	26	131	9.2	2121	42	POABETA2	P.agilis beta-2 tubul	1.14e+01
c	27	130	9.2	6504	27	AAU02548	Aedes aegypti UGALS v	1.41e+01
c	28	130	9.2	8780	31	MQSVITE	Aedes aegypti (clone	1.41e+01
c	29	131	9.2	11098	26	SYCSLRD	Synechocystis sp. slr	1.14e+01
c	30	128	9.0	1477	22	LEP1S1533	Leptospira borgpeters	2.13e+01
c	31	127	8.9	684	47	CJDNASRY	C.jacchus SRY gene.	2.61e+01
c	32	127	8.9	830	26	TRNTETE	Transposon Tn10 (from	2.61e+01
c	33	126	8.9	1337	79	XLTF3A36	Xenopus laevis gene e	3.19e+01
c	34	127	8.9	2074	41	LEPOOXE	L.esculentum gene for	2.61e+01
c	35	127	8.9	2146	43	S40548	polyphenoloxidase [ly	2.61e+01
c	36	127	8.9	2214	40	EGOCNPDHF	Erigeron hybridus chl	2.61e+01
c	37	127	8.9	2226	41	HAYCPNDHF	Haplocarpha scaposa c	2.61e+01
c	38	127	8.9	2226	39	ARZCPNDHF	Arctotis stochadifoli	2.61e+01
c	39	127	8.9	2772	18	AVINIFANFA	A.vinelandii nitrogen	2.61e+01
c	40	126	8.9	4032	53	HUM23DC1Z	Homo sapiens (subclon	3.19e+01
c	41	127	8.9	7194	50	HSENO3	H.sapiens ENO3 gene f	2.61e+01
c	42	127	8.9	10603	75	PAMFNL	Sendai virus (strain	2.61e+01
c	43	127	8.9	20235	25	SERERYAB	S.erythraea second an	2.61e+01
c	44	127	8.9	20444	25	SEERYABS	S.erythraea eryA gene	2.61e+01
c	45	126	8.9	40780	4	CEM03C11	Caenorhabditis elegans	3.19e+01

ALIGNMENTS

RESULT 1
LOCUS HSU06654 752 bp mRNA PRI 30-JUL-1994
DEFINITION Human differentiation antigen melan-A protein mRNA, complete cds.
ACCESSION U06654
NID g517022
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens

181 WSNYTCYCTGGNACNCAATGCGCTNACNMGMNCTGYCCNCAAGGNTTGYA 240

Db 174 ctgctcatcgctgttggtattgtagaagacgaatggatacagagcctgatggataaa 233

Qy	121	YTYNTNATHGGNTGYTGATYTCYMGNMGMNAAAYGNTAYMNGMGCNTYNTATGGYAAAR	180
Db	234	agcttcattcatgtgcaactcaatgtgccttaacaaagaagatgcccaagaagggttgat	293
Qy	181	WSNYTNCAYCTGCGNACNARTCYGNTVACNMGMNGTGYCCNARGCAGGNTTYGAY	240
Db	294	catcggaagcaaaagtgtcttccaagagaaaaactgtgaacctgtgttcccaatgct	353
Qy	241	CAYMGCAYSWNAAGTWSNYTNCARGAARAATYTGARGCTGTGTCNCNAAACN	300
Db	354	ccacctgttatgaaactctctcagaacagtcaccacacacctattcacc	406
Qy	301	CNCNCNGTAYGARAATYWSNGNCARGAWSNCCNCCNCONATYWSNCC	353
RESULT	3		
LOCUS	S77127	12857 bp	DNA
DEFINITION	manganese superoxide dismutase/MnSOD [human, embryonic lung fibroblast cell line W138, Genomic, 12857 nt].		
ACCESSION	S77127		
NTD	g998582		
KEYWORDS	human embryonic lung fibroblast cell line W138.		
SOURCE	Homo sapiens		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 12857)		
AUTHORS	Wan, X.S., Devalaraja, M.N. and St Clair, D.K.		PRI 26-SEP-1995
TITLE	Molecular structure and organization of the human manganese superoxide dismutase gene		
JOURNAL	DNA Cell Biol. 13 (11), 1127-1136		(1994)
MEDLINE	95217333		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 164378] from the original journal article. This sequence comes from Fig. 3.		
COMMENT	Map location: 6.		
FEATURES	NCBI gi: 998582		
source	Location/Qualifiers		
	1..12857		
	/organism="Homo sapiens"		
	/note="human"		
mRNA	join(857..881,1179..1363,5759..5876,8960..9140,11351..11498)		
	/note="Description: manganese superoxide dismutase/MnSOD"		
CDS	join(857..881,1179..1363,5759..5876,8960..9140,11351..11498)		
	/partial		
	/note="coding region is determined by the author; Description: manganese superoxide dismutase/MnSOD"		
	/codon start=1		
	/translation="MLSRVCGSGVSGIQAEAQPPRALIRPCTSHQRADHAA"		
BASE COUNT	3335 a 2627 c 3147 g 3748 t		
ORIGIN	FQQAPGRIREQPERHREVGPGVGCGMQLPR"		
Query Match	12.1%	Score 172;	DB 59; Length 12857;
Best Local Similarity	40.3%	Pred. No. 9.32e-04;	
Matches	31; Conservative	20; Mismatches 26;	Indels 0; Gaps 0;
Db	12046	agcagtgtaactgaactgggttgctcaactgcagtaagaagcgaaacttcaacgca	12105
Qy	214	MGMNGTGYCCNARGGNTTYGAYCAYMGCAYSWNAAGTWSNYTNCARGAAR	273
Db	12106	aactgcataaagggtggt	12122

QY	274	AAVTGARGCNGTNGT 290	11:11:11:11
RESULT	4		
LOCUS	ACU12386	41591 bp	DNA circular INV 24-AUG-1995
DEFINITION	Acanthamoeba castellanii mitochondrion, complete genome.		
ACCESSION	U12386		
NID	9562028		
KEYWORDS	23S-like ribosomal RNA; YMF46; YMF47; YMF48; Ile-tRNA; Ala-tRNA; Pro-tRNA; Asp-tRNA; initiator Met-tRNA; 16S-like ribosomal RNA; cytochrome oxidase subunit 1; cytochrome oxidase subunit 2; ribosomal protein S4; Phe-tRNA; X-tRNA; NADH dehydrogenase subunit 6; NADH dehydrogenase subunit 5; NADH dehydrogenase subunit 11; cytochrome oxidase subunit 3; NADH dehydrogenase subunit 4; NADH dehydrogenase subunit 2; ribosomal protein S2; His-tRNA; Leu-tRNA; H(+)-transporting ATPase subunit 9; Tyr-tRNA; cytochrome b; NADH dehydrogenase subunit 41; ORF83; ORF115; H(+)-transporting ATPase subunit 1; NADH dehydrogenase subunit 1; Trp-tRNA; ribosomal protein L11; ribosomal protein S12; ribosomal protein S7; ribosomal protein L2; ribosomal protein S19; ribosomal protein S3; ribosomal protein L16; ribosomal protein L14; ribosomal protein L5; ribosomal protein S14; ribosomal protein S8; ribosomal protein L6; ribosomal protein S13; ribosomal protein S11; ORF8; ORF25; NADH dehydrogenase subunit 3; NADH dehydrogenase subunit 9; NADH dehydrogenase subunit 7; H(+)-transporting ATPase subunit 6; ORF349; Gln-tRNA; Lys-tRNA; Glu-tRNA.		
SOURCE	amoeba.		
ORGANISM	Mitochondrion Acanthamoeba castellanii		
	Eukaryotae; mitochondrial eukaryotes; Acanthamoebidae; Acanthamoeba.		
REFERENCE	1 (bases 1 to 41591)		
AUTHORS	Burger, G., Plante, I., Loneragan, K.M. and Gray, M.W.		
TITLE	The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii. Complete sequence, gene content and genome organization		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 7352)		
AUTHORS	Loneragan, K.M. and Gray, M.W.		
TITLE	The ribosomal RNA gene region in Acanthamoeba castellanii mitochondrion DNA. A case of evolutionary transfer of introns between mitochondria and plastids?		
JOURNAL	J. Mol. Biol. 239 (4), 476-499 (1994)		
MEDLINE	94275838		
REFERENCE	3 (bases 5124 to 5757)		
AUTHORS	Loneragan, K.M. and Gray, M.W.		
TITLE	Editing of transfer RNAs in Acanthamoeba castellanii mitochondria		
JOURNAL	Science 259 (5096), 812-816 (1993)		
MEDLINE	93157849		
REFERENCE	4 (bases 1 to 7352)		
AUTHORS	Loneragan, K.M. and Gray, M.W.		
TITLE	Predicted editing of additional transfer RNAs in Acanthamoeba castellanii mitochondria		
JOURNAL	Nucleic Acids Res. 21 (18), 4402 (1993)		
MEDLINE	94021401		
REFERENCE	5 (bases 5124 to 5757)		
AUTHORS	Loneragan, K.M. and Gray, M.W.		
TITLE	Subunits 1 and 2 of cytochrome c oxidase are expressed from a continuous open reading frame in the mitochondrial DNA of Acanthamoeba castellanii		
JOURNAL	Unpublished		

REFERENCE 6 (bases 1 to 41591)

AUTHORS Burger, G.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1994) Gertraud Burger, University of Montreal, Biochemistry, 2900 Boul. Edouard-Montpetit, Montreal, Quebec, H3C 3J7, Canada

COMMENT NCBI gi: 562028

FEATURES

source

Location/Qualifiers

1..41591

/mitochondrion

/strain="Neff (ATCC 30010)"

/organism="Acanthamoeba castellanii"

join(37..1806,2562..2581,3442..4074,4828..5123)

/gene="rnl"

/note="large subunit (LSU) ribosomal RNA"

/citation=[1]

/citation=[2]

/product="23S-like ribosomal RNA"

37..1806

/gene="rnl"

/note="large subunit (LSU) ribosomal RNA"

/number=1

/citation=[1]

/citation=[2]

/product="23S-like ribosomal RNA"

1807..2561

/note="group I intron, AcLSU.m1"

/number=1

/citation=[1]

/citation=[2]

2012..2440

/gene="ymf46"

/note="rnl intron 1 (AcLSU.m1) orf, group I orf; NCBI gi:

562029"

/codon_start=1

/citation=[1]

/citation=[2]

/transl_table=4

/product="YMF46"

/db_xref="PID:g562029"

/translation="MKIDQWIVGVGEGCFYIGINKSVDSKLGQYVPEERVQHK

RDIKLYIAKDOFFGHGSVVCNKSNGSELYEYRVRKFETLHDVILPFFESNGLTSKPF

NFLAFRDVILMKREHLTSGLSKIIDIKSMNRSIH"

2562..2581

/gene="rnl"

/number=2

/citation=[1]

/citation=[2]

2582..3441

/note="group I intron, AcLSU.m2"

/number=2

/citation=[1]

/citation=[2]

2849..3355

/gene="ymf47"

/note="rnl intron 2 (AcLSU.m2) orf, group I orf; NCBI gi:

562030"

/codon_start=1

/number=2

/citation=[1]

/citation=[2]

/transl_table=4

/product="YMF47"

/db_xref="PID:g562030"

/translation="MKNLDAQKALIDPFWISGFVDGEGCFISFNKERLTGLIEV
RPFSEISQTRDKEGIALRCIADFLNFFDGGFIRFSKRDNTWKYECRDLSDIRSKVLPH
FEFTLRTKLRDFELFKDVVNSVASKQHINEVGIAKRIIDISYQINMGKKLTKDELL
SKINLNKL"

exon

3442..4074

/gene="rnl"

/number=3

4075..4827

/note="group I intron, AcLSU.m3"

/number=3

/citation=[1]

/citation=[2]

4207..4701

/gene="ymf48"

/note="rnl intron 3 (AcLSU.m3) orf, group I orf; NCBI gi:

562031"

/codon_start=1

/number=3

/citation=[1]

/citation=[2]

/transl_table=4

/product="YMF48"

/db_xref="PID:g562031"

/translation="MOKNQFKLNQLAYLAGFEADGCFIQQIPLQYRYKHTIR

ISIVFYOKKDKHWFIQLKNLIGLSIRFNDGMLEYISITGLSNKFLFMILFPYVIL

KKNLAVLIPRIIKGLNDVRNENGFLVCKLVDEVADHTYSKKRNTSLTVKNSLLLPV

ETEE"

4828..5123

/gene="rnl"

/number=4

/citation=[1]

/citation=[2]

5157..5227

/gene="trnI2"

/note="Gene originally designated trnM1 [2], [3]; now

identified as a homolog of the E. coli gene encoding an

Ile-tRNA with anticodon LAU, where L is lysidine, a

modified C"

/anticodon=(pos:unspecified,aa:Ile)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Ile-tRNA"

5258..5328

/gene="trnA"

/anticodon=(pos:unspecified,aa:Ala)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Ala-tRNA"

5336..5408

/gene="trnP"

/anticodon=(pos:unspecified,aa:Pro)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Pro-tRNA"

5412..5484

/gene="trnD"

/anticodon=(pos:unspecified,aa:Asp)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Asp-tRNA"

AUTHORS Kaiser, E., Eberhard, D. and Knippers, R.
TITLE Exons encoding the highly conserved part of human glutamyl-tRNA synthetase

FEATURES

Query Match 9.6%; Score 137; DB 50; Length 267;

```

/product="transforming growth factor-beta type III
receptor"
/db_xref="PID:g511843"
/translation="MSSSCVTPFALMILCFADAGPVRSECVLSPVNSHPHVHALL
SFTVLGSGASRGTTGLPQEVVIVLINRPDEGLDHREVEVTHLTPISVTHIQKPLVF
LIANSPLPLVWKLTERTLAPGIRRVFTVSSGIVMFEKGNFSLSAETEKNFPERNEHL
LMAOKEGYCATVSTELKTSNRYIKVGDDVFPPTKTEKNFSLNLYLAGYLAQPKRA
EGCLMNLVQRESEVHIELTPNSNYSYAFQVDIIVDINKPSQPGAKLERNVVLIICK
KSNVNIKSHDVQKGLVEVTSIGTGEKTERSMTKSVIPDIPSSHSLIKWAYEH
KYSPTS YTPKAVANRFLQLETHEMDNDEHSLPELTELHGAKNPALSGDLTFP
KYPNCGHETGGEGIFPSSRDSVDTLIHDFEHSLSKHKEPEEVQSGADVALS IKCDDP
VMTVAEKDSQASGRTVETLISLDHSCAKMNGTHF LLESLPNKQGTSTVTLDKTV
YFNS YIQLSSPAGSGFDDMDGNGCFGDAEGDVTF SNMPEAFENCT LHQEK
DFENPMFEPPEPHI NVTFENMELYTDLE LAPSQGLFSVAEGCP IYVEVSVTKADIS
FVFKPFINI SLFL LHCELTCTNIDQRTQRLPKCPVDEACTSLNVDMTI LAMNNKRTI
ETKPLVITHEKQEDSLSLPKSNVQKSVFVYGLDITLVVGIAPFAAFVIGALLTGALMFI
YSHTGAARQKQVPTSPPASNSAHSIGTQSTPCSSSNAT"
BASE COUNT 996 a 712 c 708 g 931 t
ORIGIN
Query Match
Best Local Similarity 38.1%; Pred. No. 3.21e+00;
Matches 24; Conservative 16; Mismatches 23; Indels 0; Gaps 0;
Db 346 tgcctcaatctcagaatacccgatgaaggtcttgaccatcatgaaagagaggtcacatt 405
QY 204 YGONTNACNMGNMGNTGYCCNCARGAGGNTTYGAYCMNGAYMSNAARGTWNSSYT 263
Db 406 aca 408
QY 264 NCA 266

```

	RESULT	11			
	LOCUS		BRNOTCH	7471 bp	RNA
	DEFINITION		B.rerio Notch mRNA..		VRT
	ACCESSION		X69088		07-DEC-1993
	NID		g433866		
	KEYWORDS		Notch; transmembrane protein.		
	SOURCE		zebrafish.		
	ORGANISM		Danio rerio		
			Eukaryotes; mitochondrial eukaryotes; Metazoa/Eumycota group;		
			Mecarzoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;		
			Vertebrata; Gnathostomata; Osteichthyes; Actinopterygii;		
			Actinopteri; Holstei; Halecostomi; Teleostei; Ellopocephala;		
			Clupeocephala; Euteleosteii; Ostariophysi; Otophysii; Cypriniformes;		
			Cyprinoidea; Cyprinidae; Rasborinae; Danio.		
	REFERENCE		1 (bases 1 to 7471)		
	AUTHORS		Bierkamp, C.		
	TITLE		Direct Submission		
	JOURNAL		Submitted (02-Nov-1992) to the EMBL/GenBank/DDBJ databases. C.		
			Bierkamp, Institut fuer Entwicklungsbiologie, Univ zu Koeln,		
			Gyrhofstr. 17, 5000 Koeln 41, FRG		
	REFERENCE		2 (bases 1 to 7471)		
	AUTHORS		Bierkamp, C. and Campos-Ortega, J.A.		
	TITLE		A zebrafish homologue of the Drosophila neurogenic gene Notch and its pattern of transcription during early embryogenesis		
	JOURNAL		Mech. Dev. 43 (2-3), 87-100 (1993)		
	COMMENT		94128602		
	FEATURES		NCBI gi: 433866		
			Location/Qualifiers		
			1..7471		
	source		/organism="Brachydanio rerio"		

```
/germline
/dev stage="embryonic (3-15 h)"
/clone lib="embr.cDNA"
/clone="C10.1.1, C10.2.1, C2.3.1, E4.1, E7.1"
158..217
sig_peptide

CDS
/genes="Notch"
/notes="transmembrane protein"
158..7471
/genes="Notch"
/notes="NCBI gi: 433867"
/codon start=1
/product="transmembrane protein-precursor"
/db xref="PIR:g433867"
/translation="MNRFLVKLITLTAASLATVAQQRCEYQNGGICETKPSGEAS
CRCPADFCACQFPNPNPSPCRNGVGRPMQNGEVGKDCVLGFSRLCLTPVN
HACWNSP RNNGTCSLLITDFTCRQPCGMSKTCQLADPCASNP CANGCQCSAFESH
YICTPPNPHGQTCQDVNCAVSPFCRNGGTCINEVGSYLCRCPPEYTPHQRLY
QPCLPSPCRSGGTCVSTDTHTGSCLPGLGTGQTCENHVDCTQHAGENGPCIDGIN
TYNCHCKHTGQCTEDVDECELSNPAQNGGTCNHTTGGFHCVCVNGTGDGSEN
IDDCASAGCSHGATCDHRVASFFCFCHRTGLLCHLDACISNPQKGSNDYNPVS
GKAICTCPPVGTSGACNODIECSLGNPCHEGGRCLNTKGSFCKGLACGYGRPCM
DVNEKSNPQNDATCLDQIGGFHCICMGFEGVFCQINSDDCASQPCFLNGKIDKIN
SFHCPCPGFSGSLOQVDVEACSTPCRNKACTDGNPKYTCETPGFSGHCELDIN
ECASSPCHVGRGVASTCDRCPGYTGRLCETNINECLSPCRNGGTCQDRENAVYI
CTPKGTGTGNCENIDDCRKPCKDKINGECYCEPFGYSGSKNINIDDCAL
NPNCHGGTCTDGVNFTCLCPDFRDTCLDSHNECSNPCIHGSCLDQINSYRCVE
AGWNGRNCIDINELSNPCVNGGTCMDTSCYLCTCRAGFGSPNQMINECASNPC
LNQSGCSDVAGKSNCLMPLTYGVCENVLAPCSPCRNGGVCRESDFQFSNCP
AGAQGTCEVDINECVNPNCTNGGVENLRGGFCRNPCTGALCENDIDDCENPC
SNGGVGRVNGFVVCILAGRGCAEDIDECVSPCRNGGCTDCVNSYTCSPAG
ESGNGCENIDPCTESSCNGCTVDCGISFSCYCLPFTGNYCOHDVNECDSPQCM
GGSCDGYGTCTCPHGTYGLNCQSLVRMCDSPCRNGGSCMQGASFTQCGASGT
GTYCDPVSVSEAAARQQGVSAVLCHAGQCVDAGNTHLCRCQAGTGTSCYQEQVDE
COPNCPQNCATDYLGGYSCVCPVGYHGMNCSKEINECLSPCQNGGTCIDLNTVK
CSPRGTQWHECIDDDCSPSPDLPFGEPCFNGRCVDRVGGVCGVCPAGFVGERC
EGVNECLSDPCDSGNYCVQLINDFRCERTGYTKRCETVFNCKDTPCKNGTC
AVASNTKHCYCLCKCPGSCSYCEYDSCGSLRCRNGATCVSHLSPRCIAPGFCG
HECQTRMDSPLVNPCTNGGTCQPIISDAFYRCSPPANFNLLCHILIDYSFGGGRD
IAPFVEVEIRCEIAEGEAGNNAICDQCNHACQWDGCSINFDPPQNCASALQC
WRYFDGKQDEQCATGLYDGCQRLEGCQNP LYDQYCRDHYADGCHDCGNNACG
EMGLCDADDPQKLAVGSLVLVHIDPDELNRSSSFLELSLLHTNVFRRANG
EALIFPYGSEHLSKHSRSDTDPQGLAQRRHSLTSFLAPTRRELDHMEVGSIV
YLEIDNRQCFQSDCEFSQSDVAFAFGLAASSGNLAVPYIEAVTSEGGPKTCEMY
PMFLVLLALAVLAAVGVVSVRRKREHGQLWPFEGFKVNEPKKRRPVEDGSVL
KPLKNSDLSIMDEQLSEWADDYTKRFRFEGQSI LEMSQQLDHRQMTQQLDAADJRL
NSMAPTPQOQIENDCMVNVGRPDGFTPLMIAASCSSGGLNENGEAEEDPSADVITD
FIYHGANTJNQTDRTGETALHAARYARSDAAKRLLESCADANVDNMGRTP LHAVA
ADAGQVQLILNRNDTDLARHDGFTPLILATRIAVEGMVEELINCHADPNVDDSG
KSA LHWAAVNNVDAVAVLLKNGKANDLONKEETPLFLAAREGSYETAKVLLHLAN
RDIADHLQLPDRIAHRRHHDIYVRLLEENLVRSPPPLPSPLCCPNTYLKFSFG
NNNTAKRTPKPGKSGKSDKIDTRKKRSGDGKNGGMEVGLSPVLSLEPHG
YLSVDSPPMPTSPFOQSPPIISLQIAGLADSHMGALQGLKRPDPSAPRILSLPVAN
NVGGAQAGACWILQRVQOQQOQQOQQOQAGMLPMTLSA NMPQVNGYPTMSSHLGA
PSHMTAHQNMAPPQHQSHTSHF LGLDLSGLDIQSSSGHAP IQTILPDQSBNAPP LLSST
QFLTPPSQHSYSNPMNDTNRHQVDPDHPFLTPSAGSPDQWSSSSPSHSLNLSDEGIS
PPTSMQWNIHPEAK"
mat_peptide
218..7468
/genes="Notch"
/product="transmembrane protein"
BASE COUNT 1726 a 2046 c 2104 g 1595 t
ORIGIN
```

Query Match 9.6%; Score 137; DB 77; Length 7471;
Beet Local Similarity 31.3%; Pred. No. 3.21e+00;

```
Matches 61; Conservative 36; Mismatches 97; Indels 1; Gaps 1;
Db 2646 cttgtctccagtcctgtaaaaacggcggtgtgtgtcgtgaatctgaggtattccaga 2705
:::||| | : : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 334 SYTGTCNGCNSNARYTTCRTTANGCNGNGGCGRTTNGNACNACNGGTYTCRCART 275
:::||| | : : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2706 gcttcctcgttaactgtccgctggatggcaa-ggtcaaacgtgtgaggtgacattaat 2764
:::||| | : : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 274 TTTTTCYTCGNARNWSNACYTNNWSTCNCKTGTCTCRAANCCTCTCTGNGRCANCKNC 215
:::||| | : : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2765 gagtgtgtgaggaaccgctgactaataggagagtggtgtaaaacctgcgcggtgattt 2824
:::||| | : : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 214 KNGTNARNGRCAYTCNGTNCNACRTGNARWSTYTRTCCATNARNGCCKRTANCRT 155
:::||| | : : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2825 cagtgcgctgcaat 2839
::: | : | : | | | | |
Cp 154 TNCNCKNCKRCART 140
```

```
RESULT 12
LOCUS CELF47F2 24537 bp DNA INV 25-NOV-1995
DEFINITION Caenorhabditis elegans cosmid F47F2.
ACCESSION U04943
NID g1072202
KEYWORDS
SOURCE Caenorhabditis elegans strain=Bristol N2.
ORGANISM Caenorhabditis elegans
Eukaryotes; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
Rhabditidae; Caenorhabditis.
REFERENCE
1 (bases 1 to 24537)
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favell, A.,
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N.,
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A.,
Saunders, D., Showkneen, R., Smaildon, N., Smith, A., Sonhammer, E.,
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,
Wilkinson-Sproat, J. and Wohldman, P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans
```

```
JOURNAL Nature 368 (6466), 32-38 (1994)
MEDLINE 94150718
REFERENCE
2 (bases 1 to 24537)
Bentley, D.
TITLE The sequence of C. elegans cosmid F47F2
JOURNAL Unpublished (1995)
REFERENCE
3 (bases 1 to 24537)
Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1995) Robert Waterston
COMMENT Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: rwenematode.wustl.edu and jee@sanger.ac.uk
NEIGHBORING COSMID INFORMATION:
5' cosmid is 2C64, 200 bp overlap; 3' cosmid is T22B2. Actual start
of this cosmid is at base position 29748 of CELF2C64.
```

221
:1
221

DEFINITION

DEFINITION

BASE COUNT	693 a	947 c	966 q	681 t
------------	-------	-------	-------	-------

RESULT	15	ONTCVG6	708 bp	RNA	MAM	15-FEB-1993
LOCUS						
DEFINITION		O.aries rearranged T-cell receptor gamma.				

```

JOURNAL      EMBO J. 12, 712-724 (1993)
COMMENT      and developmentally regulated expression of the sheep gamma/delta T
FEATURES
source       cell receptor
             NCB1 gi: 2278
             Location/Qualifiers
               1..708
               /organism="Ovis aries"
               /strain="White Alpine"
               /dev_stage="61 day old fetus"
               /cell_type="peripheral lymphocytes"
             79..>708
             /note="NCBI gi: 2279"
             /codon_start=1
             /product="TCR gamma"
             /db_xref="PID:g2279"
             /translations="MGFSLGAGEGRAVLGRLLALLWLVPVGIOQFIRLSQSRVWVGSAG
             GATWPCRHVSRSVDYVHFRLEGQAPPELLNLALSKRDVQVDSVLGGCKVKSAAARGD
             GKSTCSLRKLSKDEGLYCAAMDYNLYNYMIFGGTKVFQDKNLPTDIIKPKRTIFL
             PSINEVNHQQTATYLCLENEFPDVIKVSKWEKNGKRVLPQSHGNTMKTN"
             79..147
             148..>708
             /product="TCR gamma"
             148..510
             /note="TCR V gamma 6 +J"
             168 a 188 c 191 g 161 t
BASE COUNT  168 a 188 c 191 g 161 t
ORIGIN

```

Query Match 9.3%; Score 132; DB 34; Length 708;
Best Local Similarity 37.8%; Pred. No. 9.28e+00;
Matches 28; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

Db 604 tgtctctggaaaatttttccctgacgttatttaaggtttcttgaaagagaagaatggc 663
||:| |:|: ||: |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

Db	664	aaaagagttctgcc	677
		:	
Oy	280	GARCCNGTNGTNCC	293

Search completed: Tue Aug 27 18:41:38 1996
Job time : 276 secs.

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 08:08:49 1996; MasPar time 3.69 Seconds
282.669 Million cell updates/sec

Tabular output not generated.

Title: >US-08-231-565A-2

Description: (1-118) from US08231565A.ppt

Perfect Score:

Sequence: I MPREDAHFIYGYPPKKGHGS.....NAPPAYEKLSAEQSPPPFYSP I18

scoring table. PAM 150

Gap 11

Searched: 76839 seqs, 883355 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq23

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13 14:part14 15:part15

Statistics: Mean 29.335; Variance 113.395; scale 0.259

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
1	889	100.0	118 15	R84212	MART-1 melanoma anti g	1.71e-83
2	889	100.0	118 11	R63158	Tumour rejection anti l	1.71e-83
3	90	10.1	380 1	R05433	cPA-P2 Hybrid plasmin	3.71e+00
4	83	9.3	200 15	R82900	Mouse B7-1	1.37e+01
5	83	9.3	212 15	R82902	Mouse B7-1 IgV-like i	1.37e+01
6	83	9.3	306 15	R82893	Mouse B7-1 alternativ	1.37e+01
7	83	9.3	306 12	R67990	Murine B lymphocyte a	1.37e+01
8	82	9.2	897 4	R20982	Sequence of beta-chai	1.65e+01
9	81	9.1	455 5	R28757	Hepatocyte nuclear fa	1.98e+01
10	81	9.1	855 10	R51252	FIV UK2 envelope prot	1.98e+01
11	81	9.1	856 10	R51248	FIV petalumna envelope	1.98e+01
12	81	9.1	856 10	R51249	FIV PPT-F14 envelope	1.98e+01

13	81	9.1	856	10	R51247	FIV envelope protein	1.98e+01
14	79	8.9	234	3	R13844	Fusaric acid resistant	2.84e+01
15	79	8.9	854	10	R51250	FIV PPR envelope prot	2.84e+01
16	79	8.9	856	10	R51254	FIV DUTCH19K1 envelop	2.84e+01
17	79	8.9	940	2	R07070	Fibronectin-binding p	2.84e+01
18	76	8.5	524	2	R05077	Sali restriction prod	4.86e+01
19	76	8.5	524	1	P90113	Rabies G protein	4.86e+01
20	76	8.5	854	8	R44957	Feline Immunodeficien	4.86e+01
21	76	8.5	1822	10	R55273	Beta eubunit of integ	4.86e+01
22	75	8.4	117	2	R07997	Human CD2 cytoplasmic	5.80e+01
23	75	8.4	342	13	R74221	Epitope on the primar	5.80e+01
24	75	8.4	335	15	R77026	Canine herpesvirus gD	5.80e+01
25	75	8.4	351	4	R20803	Human CD2 antigen.	5.80e+01
26	75	8.4	351	2	R11921	Human TII sugar prote	5.80e+01
27	75	8.4	353	2	R06365	Soluble two domain hu	5.80e+01
28	75	8.4	360	1	P81178	Sequence of human TII	5.80e+01
29	75	8.4	855	10	R58585	FIV UK8 envelope prot	5.80e+01
30	75	8.4	855	10	R51255	FIV SWISS22 envelope	5.80e+01
31	75	8.4	870	9	R51246	Composite sequence of	5.80e+01
32	74	8.3	132	7	R36393	Penetrin.	6.91e+01
33	74	8.3	577	1	R04924	Immunoprotein TANG12.	6.91e+01
34	74	8.3	1026	11	R58906	Human protocaderhin-4	6.91e+01
35	74	8.3	1203	11	R58911	Product of alternativ	6.91e+01
36	73	8.2	10	15	R84778	MART-1 melanoma anti	8.23e+01
37	73	8.2	238	2	P70375	Human interleukin-2 r	8.23e+01
38	73	8.2	272	3	P50001	Sequence of interleuk	8.23e+01
39	73	8.2	272	14	R79389	Human interleukin-2 r	8.23e+01
40	73	8.2	272	3	P61052	Sequence encoding int	8.23e+01
41	73	8.2	896	4	R21580	Plasma membrane proto	8.23e+01
42	72	8.1	10	15	R84777	MART-1 melanoma anti	9.80e+01
43	72	8.1	479	1	P81013	Complete sequence of	9.80e+01
44	72	8.1	479	3	P98034	Sequence encoded by t	9.80e+01
45	72	8.1	1091	2	P98500	Partial sequence enco	9.80e+01

ALIGNMENTS

RESULT	1
ID	R84212 standard; Protein; 118 AA.
AC	R84212;
DT	20-APR-1996 (first entry)
DE	MART-1 melanoma antigen.
KW	MART-1; melanoma antigen recognised by T-cell; melanoma;
KKW	metastatic melanoma; tumour-associated antigen; immunogen;
OS	diagnosis; prognosis; prophylaxis; therapy; vaccine.
QS	Mammalian.
Key	Location/Qualifiers
Region	Z1..47
FT	/note= "hydrophobic region"
FN	W09529193-A2.
PPD	02-NOV-1995.
PPF	21-APR-1995; U05063.
PPR	22-APR-1994; US-231565.
PPRP	05-APR-1995; US-417174.
PPPP	(USSH) US SEC DEPT HEALTH.
PPI	Kawakami Y, Rosenberg SA;
PPII	WPI; 95-382963/49.
DR	N-PSDB; T02714.
DRR	
PPT	DNA encoding melanoma antigens recognised by T-lymphocytes - also
PPTT	vectors, host cells and antibodies, used to detect, treat and
PPTTT	immunise animal against melanoma.
PPTTTT	Claim 11; Page 117; 184pp; English.
PPS	The melanoma antigen (MART-1) is produced by recombinant DNA
PPCC	methods, i.e. preferably using a baculovirus vector for expression

CCC in insect cell cultures. MART-1 protein is a source of immunogenic peptides (see R474196 for peptide M9-2) which are optionally modified (see R847783-R849800) and used in medicaments for the treatment or prevention (by immunization) of melanoma. Antibodies against MART-1 and its immunogenic peptides may be used in the detection and isolation of MART-1 from a sample, the detection of which is indicative of a disease state (melanoma or metastatic melanoma). Sequence 118 AA; CSQ

Query Match	100.0%;	Score 889;	DB 15;	Length 118;
Best Local Similarity	100.0%;	Pred. No. 1.71e-83;		
Matches	118;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Ddb	1	mpredahfiygyppkghghsyttaeaaagilktvlglwlllgccwrrrrgyralmdk	60	
Qy	1	MPREDAHFIYGYPKGHGHSYTTAEAAAGILGTVLGLWLLLGCCWRRRRGYRALMDK	60	
Db	61	slhvtgcaltrrcpqegfhdrrskvslqekncpvpnpapayeklaeqspppyvp	118	
Qy	61	SLHWGTCAALTARCPQEGFHDRRSKVSLQEKNCPEVPNPAPAYEKLSAQSPPYSP	118	

RESULT	2	
ID	R63158	standard; Protein; 118 AA.
AC	R63158;	
DDT	26-MAY-1995	(first entry)
DDT	Tumour rejection antigen precursor.	
DE	Tumour rejection antigen; precursor;	H1A-A2 molecule; tyrosinase;
KW	isolation; melanoma;	cell line; LB-39-MEL; diagnosis; vaccine;
KW	therapy.	
OS	Homo sapiens.	
PN	W09421126-A.	
PD	29-SEP-1994.	
PF	09-MAR-1994;	U02487.
PF	18-MAR-1993;	US-032978.
PR	(JUDW-) LUDWIG INST CANCER RES.	
PI	Boon-Falleur T, Brichard V, De Plaen E, Traversari C;	
PI	Van Pel A, Wolfel T;	
PI	WPI; 94-316544/39.	
DR	N-PSDB; Q76370.	
TT	Nucleic acid coding for a tumour rejection antigen precursor - is	
TT	used for developing prods. for diagnosis or treatment of expression	
TT	related disorders, partic. melanoma	
PS	Claim 5; Page 14; 26pp; English.	
CCC	This sequence represents the tumour rejection antigen precursor which is	
CCC	processed to a tumour rejection antigen presented by H1A-A2 molecules.	
CCC	The tumour rejection antigen is not related to tyrosinase. The cdNA	
CCC	encoding this sequence was isolated from the melanoma cell line,	
CCC	LB-39-MEL. The tumour rejection antigen may be used for diagnosis or	
CCC	in vaccines or for therapy of disorders characterised by the expression	
CCC	of the tumour rejection antigen precursor, particularly melanoma.	
CSQ	Sequence	118 AA;

[illegible]

RESULT	3	
ID	R05433	standard; protein; 380 AA.
AC	R05433;	
DT	30-JUL-1990	(first entry)
DE	cPA-P2 Hybrid plasminogen activator.	
KW	plasminogen activator; fibrin; urokinase; thromboembolic disease; ds.	
OS	Synthetic.	
EH	Key	Location/Qualifiers
FT	Domain	1..23
FT	/label=Secretory leader.	
FT	Domain	24..102
FT	/label=Plasminogen Kringle 1	
FT	Domain	103..116
FT	/label=Urokinase linker	
FT	Domain	117..380
FT	/label=ScuPA protease domain.	
FT	W09001332-A.	
PD	22-FEB-1990.	
PF	10-AUG-1988; 02771.	
PR	10-AUG-1988; W0-U02771.	
PA	(CETU) Cetus Corp.	
PI	Halluin AP;	
DR	WPI; 90-083374/11.	
DR	N-PSDB; Q02301.	
PT	Compens. cong. plasminogen activator conjugated to heparin component -	
PT	used for treatment of thromboembolic disease, with longer half	
PT	life and improved targeting.	
PT	Disclosure; p; English.	
PS	Gene encodes hybrid plasminogen activator (PA) comprising Kringle 1, an	
CC	urokinase linker, and an urokinase protease domain wherein glycine residue	
CC	at position 158 is replaced with a lysine.	
CC	The compound is used to treat thromboembolic disease esp. with myocardial	
CC	infarction, has a longer half-life than free PA and targets the heparin	
CC	site of thrombus or embolism reducing the risk of reocclusion.	
SQ	Sequence	380 AA;

Query Match	10.1%;	Score 90;	DB 1;	Length 380;
Best Local Similarity	30.9%;	Pred. No. 3,71e+00;		
Matches	21;	Mismatches 15;	Indels 8;	Gaps 7;
Conservative	24;			
Db	7	la lllllllpgcwasectgdknyrgtmaktn-gitcckwastetprfrfspathps	65	
Qy	33	LTVILGVLLIGW-Y-CRRNRG--YRALMDKSLHVGTCAL-TRRPCEG-FDH-RDSK	85	
Db	66	egleenyc	73	
Ov	86	VSTLOKNC	93	

RESULT	4
ID	R82900 standard; Protein; 200 AA.
AC	R82900;
DT	07-MAY-1996 (first entry)
DE	Mouse B7-1 (IgV-like domain deleted).
KW	T-cell costimulatory molecule; B7-1; T-lymphocyte; CD28; CTLA4;
KW	receptor; immunoglobulin.
OS	Mus musculus.
FH	Key
FT	Location/Qualifiers
FT	Peptide 1..37
FT	/label= Sig_peptide
FT	Domain 169..200

FT /label= Cytoplasmic domain
 FT /note= "cytoplasmic domain is encoded by exon 5
 FT of the B7-1 gene"
 PN W09523859-A2.
 PD 08-SEP-1995.
 PF 02-MAR-1995; U02576.
 PR 02-MAR-1994; US-205697.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PA (DAND) DANA FARBEN CANCER INST.
 PI Borriello F, Freeman GJ, Nadler LM, Sharpe AH;
 DR WPI; 95-320574/41.
 DR N-PSDB; T01047.
 PT Novel T cell co-stimulatory molecules - corresponding to naturally
 PT occurring alternatively spliced forms of T cells co-stimulatory
 PT molecules or variants
 PS Disclosure; Page 55-56; 111pp; English.
 CC A naturally occurring form of mouse T-cell costimulatory molecule
 CC B7-1 (R82900) has the signal peptide directly linked to the
 CC IgG-like domain, i.e. the IgV-like domain is deleted. It is
 CC encoded by exons 1, 3, 4 and 5 (see T01047) of the B7-1 gene. An
 CC alternatively spliced form of IgV-deleted B7-1 (R82901) is encoded
 CC by exons 1, 3, 4 and 6. T-cell costimulatory molecules can be
 CC produced in which the IgV-like domain is deleted.
 SQ Sequence 200 AA;

Query Match 9.3%; Score 83; DB 15; Length 200;
 Best Local Similarity 27.6%; Pred. No. 1.37e+01;
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 146 gagfgavttvvviiiko-fckhrscfr 173
 :||:| : : :||:| : :||:| :|
 Qy 27 AAGIGILTIVILGVLILLGCMYCRNRNGYR 55

RESULT 5
 ID R82902 standard; Protein; 212 AA.
 AC R82902;
 DT 07-MAY-1996 (first entry)
 DE Mouse B7-1 IgV-like isoform.
 KW T-cell costimulatory molecule; B7-1; T-lymphocyte; CD28; CTLA4;
 KW receptor; immunoglobulin; interleukin-2.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT Peptide 1..37
 FT /label= Sig_peptide
 FT W09523859-A2.
 PD 08-SEP-1995.
 PF 02-MAR-1995; U02576.
 PR 02-MAR-1994; US-205697.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PA (DAND) DANA FARBEN CANCER INST.
 PI Borriello F, Freeman GJ, Nadler LM, Sharpe AH;
 DR WPI; 95-320574/41.
 DR N-PSDB; T01040.
 PT Novel T cell co-stimulatory molecules - corresponding to naturally
 PT occurring alternatively spliced forms of T cells co-stimulatory
 PT molecules or variants
 PS Disclosure; Page 49-51; 111pp; English.
 CC An alternatively spliced form of murine T-cell costimulatory
 CC molecule B7-1 (R82893) is encoded by exons 1-6 (T01040) of the B1-7
 CC gene. It includes cytoplasmic domains (R82887 and R82888) encoded
 CC by exons 5 (T01035) and 6 (T01034). Novel cytoplasmic forms of
 CC T-cell costimulatory molecules have been produced that bind to
 CC CD28 or CTLA4 and trigger a costimulatory signal in T-cells.
 SQ Sequence 306 AA;

Query Match 9.3%; Score 83; DB 15; Length 306;
 Best Local Similarity 27.6%; Pred. No. 1.37e+01;
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 252 gagfgavttvvviiiko-fckhrscfr.279
 :||:| : : :||:| : :||:| :|
 Qy 27 AAGIGILTIVILGVLILLGCMYCRNRNGYR 55

RESULT 7
 ID R67990 standard; Protein; 306 AA.
 AC R67990;
 DT 21-AUG-1995 (first entry)
 DE Murine B lymphocyte antigen B7 (mB7).
 KW B lymphocyte activation antigen; B7-1; Ig superfamily; CD28;
 KW transmembrane protein.

CC transmembrane domain, i.e. the IgC-like domain is deleted. It is
 CC encoded by exons 1, 2, 4 and 5 of the B7-1 (T01049) gene.
 CC This IgV-like isoform of B701 was expressed in CHO cells. It
 CC triggered a costimulatory signal in T-cells, causing stimulation
 CC of interleukin-2 prodn.
 SQ Sequence 212 AA;

Query Match 9.3%; Score 83; DB 15; Length 212;
 Best Local Similarity 27.6%; Pred. No. 1.37e+01;
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 158 gagfgavttvvviiiko-fckhrscfr 185
 :||:| : : :||:| : :||:| :|
 Qy 27 AAGIGILTIVILGVLILLGCMYCRNRNGYR 55

RESULT 6
 ID R82893 standard; Protein; 306 AA.
 AC R82893;
 DT 07-MAY-1996 (first entry)
 DE Mouse B7-1 alternatively spliced form.
 KW T-cell costimulatory molecule; B7-1; T-lymphocyte; CD28; CTLA4;
 KW receptor; immunoglobulin.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT Peptide 1..37
 FT /label= Sig_peptide
 FT W09523859-A2.
 PD 08-SEP-1995.
 PF 02-MAR-1995; U02576.
 PR 02-MAR-1994; US-205697.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PA (DAND) DANA FARBEN CANCER INST.
 PI Borriello F, Freeman GJ, Nadler LM, Sharpe AH;
 DR WPI; 95-320574/41.
 DR N-PSDB; T01040.
 PT Novel T cell co-stimulatory molecules - corresponding to naturally
 PT occurring alternatively spliced forms of T cells co-stimulatory
 PT molecules or variants
 PS Disclosure; Page 49-51; 111pp; English.
 CC An alternatively spliced form of murine T-cell costimulatory
 CC molecule B7-1 (R82893) is encoded by exons 1-6 (T01040) of the B1-7
 CC gene. It includes cytoplasmic domains (R82887 and R82888) encoded
 CC by exons 5 (T01035) and 6 (T01034). Novel cytoplasmic forms of
 CC T-cell costimulatory molecules have been produced that bind to
 CC CD28 or CTLA4 and trigger a costimulatory signal in T-cells.
 SQ Sequence 306 AA;

Query Match 9.3%; Score 83; DB 15; Length 306;
 Best Local Similarity 27.6%; Pred. No. 1.37e+01;
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 252 gagfgavttvvviiiko-fckhrscfr.279
 :||:| : : :||:| : :||:| :|
 Qy 27 AAGIGILTIVILGVLILLGCMYCRNRNGYR 55

RESULT 7
 ID R67990 standard; Protein; 306 AA.
 AC R67990;
 DT 21-AUG-1995 (first entry)
 DE Murine B lymphocyte antigen B7 (mB7).
 KW B lymphocyte activation antigen; B7-1; Ig superfamily; CD28;
 KW transmembrane protein.

Mus musculus.	Location/Qualifiers
Key	Protein 1..37
/label= signal sequence	
/note= "hydrophobic"	
Domain	38..247
/label= extracellular	
/note= "6"	
Domain	248..272
/label= transmembrane	
Domain	273..306
/label= intracellular (cytoplasmic)	
Domain	38..142
/label= Ig V-set domain	
Domain	143..236
/label= Ig c-set domain	
Misc difference 1..306	
/label= published	
/note= "Freeman, G.J. et al. see CC"	
W09503408-A.	
02-FEB-1995.	
26-JUL-1994; U08423.	
26-JUL-1993; US-101624.	
19-AUG-1993; US-109393.	
03-NOV-1993; US-147773.	
(DAND) DANA FARBER CANCER INST INC.	
(REPK) REPLICEN CORP.	
Freeman GJ, Gray GS, Greenfield E, Nadler LM;	
WPI; 95-075236/10.	
N-PSDB; Q81372.	
Nucleic acids encoding CTLA4/CD28 counter receptor, B7-2 - useful	
for enhancing or suppressing T-cell mediated immune responses	
Disclosure; pages 118-120; 175pp; English.	
Q81371 is in pCDM8 vector. It is derived from germline B	
lymphocytes, cell lines 702 and A20, clones B7 nos. 1 and 29.	
It can be found in Genbank at Accession no. X60958. The encoded	
protein, R67990, binds both human CTLA4 and human CD28. It is	
related to human hB7-2 (see Q81351) and human hB7-1 (see Q81371). It	
part of R67990 (see CC) is published in Freeman, G.J. et al.	
J. of Experimental Medicine, in press at the time when the patent	
application was written.	
Sequence 306 AA;	
Sequence 306 AA;	

```
Query Match          9.3%; Score 83; DB 12; Length 306;
Best Local Similarity 27.6%; Pred. No. 1.37e+01;
Matches      8; Conservative    14; Mismatches   6; Indels   1; Gaps   1;
```

RESULT	8	
ID	R20982 standard; Protein; 897 AA.	
AC	R20982;	
DDT	17-MAY-1992 (first entry)	
DE	Sequence of beta-chain of a human granulocyte-macrophage	
DE	colony stimulating factor (GM-CSF) receptor.	
KW	Agonist; antagonist; myeloid leukaemia; therapy; screening;	
KW	diagnosis; granulocyte-macrophage colony stimulating factor.	
OS	Homo sapiens.	
Key	Location/Qualifiers	
FT	Peptide	1..17
FT	/label= signal	

FT Modified -site 58..60
FT /label= glycosylation
FT Modified -site 191..193
FT /label= as above
FT Modified -site 346..348
FT /label= as above
PN W09201788-A.
PD 06-FEB-1992.
PF 16-JUL-1991; U04846.
PR 18-JUL-1990; US-554745.
PP (SCHE) SCHERING CORP.
PI Havashida K, Kitamura T, Miyajima A;
DR WPI; 92-064947/08.
DR N-PSDB; Q21453.
DR BETA-chain of human granulocyte-macrophage CSF receptor - used
PT for screening agonists and antagonists of human GM-CSF, e.g. for
PT screening myeloid leukaemia
PT Claiming myeloid leukaemia
PS Claim 2; Page 15-18; 26pp; English.
CC The nucleic acid encoding the human GM-CSF beta-chain is isolated
CC from a cDNA library prepd. from poly(A) + RNA from TF-1 cells. The
CC high affinity human GM-CSF receptor (Kd < 1 nM) can be used for
CC screening candidate GM-CSF agonists and antagonists e.g. for
CC treating myeloid leukaemias.
SQ Sequence 897 AA;

Query Match 9.2%; Score 82; DB 4; Length 897;
Best Local Similarity 33.3%; Pred. No. 1.65e+01;
Matches 14; Conservative 14; Mismatches 10; Indels 4; Gaps 4;

Db 528 gfdsesvpltiiedpkhvcppsgpdttpaasdlpteqppsp 569
|| : { :::: : | : | :: | | ::||:| :|
ov 78 GFDRD-SKVSIQF-KN-CFQW-NAPPAYEKLSAFOSPPP 115

RESULT	9	
ID	R28757 standard; Protein; 455 AA.	
AC	R28757;	
DT	14-JAN-1993 (first entry)	
DE	Hepatocyte nuclear factor 4. (HNF4)	
KW	Hepatocyte nuclear factor 4; apolipoproteins; fat; cholesterol;	
KW	Apo CIII; Apo AI; Apo B; pyruvate kinase; alpha 1 antitrypsin;	
KW	glutamine synthetase; coronary heart disease hyperlipidaemia;	
KW	liver disease; arteriosclerosis; obesity; <i>ss</i> .	
PN	W09211365-A.	
PD	09-JUL-1992.	
PF	23-DEC-1991; U09733.	
PP	21-DEC-1990; US-631720.	
PA	(UYRQ) UNIV ROCKEFELLER.	
PI	Darnell JE, Sladek EM, Zhong W;	
PI	WPI; 92-250087/30.	
DR	N-PSDB; Q31765.	
PT	Hepatocyte nuclear factor 4 and its DNA, regulation and	
PT	antibodies - useful for treating cardiovascular diseases e.g.	
PT	atherosclerotic heart disease, hyperlipidaemia and	
PT	arteriosclerosis; also as an anorectic	
PS	Disclosure; Fig 3; 100pp; English.	
CC	This sequence was deduced from the cDNA sequence. The protein has	
CC	a structure analogous to that of steroid/hormone receptors. It	
CC	contains a region with two potential zinc fingers between amino	
CC	acids 50 and 116, which is 40 to 63% identical to the zinc finger	
CC	(DNA binding) domain of other members of the steroid receptor	
CC	superfamily. The proposed regulatory protein for mouse MHC I	
CC	(M2-R11P) had the greatest similarity (62.7%), with human thyroid	
CC	hormone receptor (c-erbA, T3-T8) having 59.7% identity in this	

1 OF RETINE IMMUNODEFICIENCY INFECTIOIN

CC Petaluma Isolates, a CC Variations in the same

CC Petaluma Isolates, a CC Variations in the same

Aug 27 07:56

US-08-231-565A-2.rag

11

CC serotypes, isolates of different geographical origin or even between
CC different isolates from the same host.
SQ Sequence 856 AA;

Query Match 9.1%; Score 81; DB 10; Length 856;
Best Local Similarity 73.3%; Pred. No. 1.98e+01;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 790 gilgiglgvlllllc 804
||| : ||||| I
Qy 31 GILTVILGVLIIIGC 45

RESULT 12
ID R51249 standard; peptide; 856 AA.
AC R51249;
DT 08-OCT-1994 (first entry)
DE FIV PET-F14 envelope protein sequence.
KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;
KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.
OS Synthetic.

FH Key Location/Qualifiers
FT Peptide 1..178
FT /note= "Hydrophobic leader peptide"
FT Region 51..66
FT /note= "V1 region"
FT Region 95..173
FT /note= "V2 region"
FT Protein 179..610
FT /note= "Surface glycoprotein"
FT Region 360..424
FT /note= "V3 region"
FT Region 451..483
FT /note= "V4 region"
FT Region 539..567
FT /note= "V5 region"
FT Protein 611..856
FT /note= "Transmembrane protein"
PN W09406471-A.
PD 31-MAR-1994.
PR 20-SEP-1993; G01974.
PR 21-SEP-1992; GB-019936.
PA (PITM) PITMAN MOORE INC.
PI Francis MJ;
DR WPI; 94-118168/14.
PT Feline immunodeficiency virus antigenic polypeptide(s) and
PT nucleic acid - used to prepare prods. for combating or diagnosis
PT of feline immunodeficiency infection
PS Disclosure; Figure 2; 68pp; English.
CC The sequences given in R51248-R51262 and R58584-85 represent the
CC feline immunodeficiency virus (FIV) envelope protein from different
CC strains and serotypes. The consensus sequence based on these, is given
CC in R51247. The synthetic FIV peptides of the invention were derived
CC principally from a combination of the sequences of the FIV UK 8 and
CC Petaluma isolates, a composite sequence of which is given in R51246.
CC Variations in the sequence may occur between different strains or
CC serotypes, isolates of different geographical origin or even between
CC different isolates from the same host.
SQ Sequence 856 AA;

Query Match 9.1%; Score 81; DB 10; Length 856;
Best Local Similarity 73.3%; Pred. No. 1.98e+01;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Aug 27 07:56

US-08-231-565A-2.rag

12

Db 790 gilgiglgvlllllc 804
||| : ||||| I
Qy 31 GILTVILGVLIIIGC 45

RESULT 13
ID R51247 standard; peptide; 856 AA.
AC R51247;
DT 08-OCT-1994 (first entry)
DE FIV envelope protein consensus sequence.
KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;
KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.
OS Synthetic.

FH Key Location/Qualifiers
FT Peptide 1..178
FT /note= "Hydrophobic leader peptide"
FT Region 51..66
FT /note= "V1 region"
FT Region 95..173
FT /note= "V2 region"
FT Protein 179..610
FT /note= "Surface glycoprotein"
FT Region 360..424
FT /note= "V3 region"
FT Region 451..483
FT /note= "V4 region"
FT Region 539..567
FT /note= "V5 region"
FT Protein 611..856
FT /note= "Transmembrane protein"
PN W09406471-A.
PD 31-MAR-1994.
PR 20-SEP-1993; G01974.
PR 21-SEP-1992; GB-019936.
PA (PITM) PITMAN MOORE INC.
PI Francis MJ;
DR WPI; 94-118168/14.
PT Feline immunodeficiency virus antigenic polypeptide(s) and
PT nucleic acid - used to prepare prods. for combating or diagnosis
PT of feline immunodeficiency infection
PS Disclosure; Figure 2; 68pp; English.
CC The sequences given in R51248-R51262 and R58584-85 represent the
CC feline immunodeficiency virus (FIV) envelope protein from different
CC strains and serotypes. The consensus sequence based on these, is given
CC in R51247. The synthetic FIV peptides of the invention were derived
CC principally from a combination of the sequences of the FIV UK 8 and
CC Petaluma isolates, a composite sequence of which is given in R51246.
CC Variations in the sequence may occur between different strains or
CC serotypes, isolates of different geographical origin or even between
CC different isolates from the same host.
SQ Sequence 856 AA;

Query Match 9.1%; Score 81; DB 10; Length 856;
Best Local Similarity 73.3%; Pred. No. 1.98e+01;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 790 gilgiglgvlllllc 804
||| : ||||| I
Qy 31 GILTVILGVLIIIGC 45

RESULT 14
ID R13844 standard; Protein; 234 AA.
AC R13844;

QY 1 MPREDAHFYGYPKKGHGHSYTTAEFAAGIGILT^{TV}VILGVILLIGCWYCRRRNGYRALMDK 60

Db 61 slhgtqcaltrrcpqegfhdhdsksvlqekncepvpvnpappayeklsaeqspypsp 118
=====

Qv 61 SLHVGTCALTRRCPQEGFHDHDSKSVLQEKNCPEVPVNPAPPAYEKLAEQSPYPSP 118
=====

2	RESULT	
ENTRY		
TITLE	S01347	#type complete
ALTERNATE NAMES	T-cell surface glycoprotein CD2 precursor - mouse	
ORGANISM	CD2 antigen; TII protein	
ORGANISM	#normal_name Mus musculus	#common_name house mouse
DATE	30-Jun-1989	#sequence_revision 30-Jun-1989
	09-Sep-1994	#text_change

ACCESSIONS
 REFERENCE
 #authors Clayton, L.K.; Sayre, P.H.; Novotny, J.; Reinherz, E.L.
 #journal Eur. J. Immunol. (1987) 17:1367-1370
 #title Murine and human T11 (CD2) cDNA sequences suggest a common signal transduction mechanism.
 #cross-references MUID:88004738

#accession S01347
##molecule_type mRNA
##residues 1-344 **##label** CIA
##cross-references EMBL:X06143
CLASSIFICATION superfamily T-cell surface glycoprotein CD2
KEYWORDS glycoprotein; transmembrane protein

```

FEATURE
1-22      #domain signal sequence #label SIG
23-344    #product T-cell surface glycoprotein CD2 #label MA
23-203     #domain extracellular #label EXT\
204-228    #domain transmembrane #label TM\
229-344    #domain intracellular #label INT
SUMMARY
#length 344 #molecular-weight 38400 #checksum 4741

```

Query Match	11.1%;	Score 99;	DB 5;	Length 344;
Best Local Similarity	40.4%;	Pred. No. 4.15e-02;		
Matches	21;	Conservative	9;	Mismatches 18;
			Indels	4;
			Gaps	4;

Db 198 pekglsf-yvtvgvqag-glllvll-valfifc-icrkrrmrkdeelei 245
||| : ||| : ||| : ||| : ||| : ||| :
Qy 13 PKKGHGHSYTTAEAGIGLTVILGVLLLCGYCRRRNGRYALMDKSLHV 64

```

3
RESULT
ENTRY
TITLE
ALTERNATE NAMES
ORGANISM
DATE
S02293      #type complete
T-cell surface glycoprotein CD2 precursor - mouse
T-lymphocyte antigen CD2
#formal_name Mus musculus #common_name house mouse
18-Dec-1989 #sequence_revision 01-Dec-1989 #text_change
18-Jun-1993

```

502293	Sewell, W.A.; Brown, M.H.; Owen, M.J.; Fink, P.J.; Kozak
502293	C.A.; Crumpton, M.J.
	Eur. J. Immunol. 1987 17:1015-1020
	The murine homologue of the T lymphocyte CD2 antigen:
	molecular cloning, chromosome assignment and cell surface
	expression.

```
#cross-references MUID:87276135
#accession S02293
##status not compared with conceptual translation
##molecule_type mRNA
##residues 1-344 ##label SEW
```

##cross-references EMBL:Y00023

#map_position 3
CLASSIFICATION #superfamily T-cell surface glycoprotein CD2
KEYWORDS glycoprotein; T-cell

```

1-22      #domain signal sequence #label SIG\
23-344    #product T-cell surface glycoprotein CD2 #label MAT
SUMMARY   #length 344 #molecular-weight 38414 #checksum 4681

```

Query Match	11.1%;	Score 99;	DB 5;	Length 344;
Best Local Similarity	40.4%;	Pred. No. 4,15e-02;		
Matches	21;	Conservative	9;	Mismatches 18; Indels 4; Gaps 4;

```
.198 peklsf-yvtvgvgag-glllvll-valfifc-ickrkrrrrkdeelei 245
      ||| : | : ||||| : | : | : ||| : | : | :
Ow   13 PKKGHGHSYTTAFEEAAGIGITVILGVLLIGCWYCRRRNGVRALMDKSLHV 64
```

RESULT	4
ENTRY	B28967
TITLE	T-cell surface glycoprotein T11 (CD2) - mouse
ORGANISM	formal name Mus musculus common name house mouse
DATE	30-Jun-1989 #sequence revision 03-Jun-1993 #text change 31-Dec-1993

ACCESSIONS	REFERENCE
B28967	Diamond, D.J.; Clayton, L.K.; Sayre, P.H.; Reinherz, E.L.
A28967	Proc. Natl. Acad. Sci. U.S.A. (1988) 85:1615-1619
#authors	Exon-intron organization and sequence comparison of human and
#journal	murine T11 (CD2) genes.
#title	

```
#cross-references MUID:88144486
#accession P28967
#molecule_type mRNA
#residues 1-344 ##label DIA
#cross-references GB:M19807; GB:J03623
#note the authors translated the codon TAT for residue 99 as
      nt.
```

```

CLASSIFICATION
KEYWORDS      #superfamily T-cell surface glycoprotein CD2
               glycoprotein
SUMMARY       #length 344 #molecular-weight 38325 #checksum 4974

Query Match    11.1%; Score 99; DB 5; Length 344;
Best Local Similarity 40.4%; Pred No. 4 15a-02;

```

```

      198 pekglaf-yvtvvgag-qllvll-valfifc-ickrkmrrrkdeeei 245
      ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Qy 13 PKKGHGHSYTTAEARAGIGLTWLVGVLLICVYCRRRNGVRLMDKSLHV 64

```

RESULT	5
ENTRY	S47272
TITLE	B-CAM protein - human
ORGANISM	formal name Homo sapiens common name man
DATE	06-Jan-1995
	#sequence_revision 06-Jan-1995
	#text_change 06-Jan-1995

ACCESSIONS
 547272
 REFERENCE
 #authors Campbell, I.G.
 #submission submitted to the EMBL Data Library, July 1994
 #accession 547272
 #status preliminary
 #molecule type mRNA

```

##residues 1-588 ##label CAM
##cross-references EMBL:X80026
SUMMARY #length 588 #molecular-weight 63566 #checksum 255

Query Match 10.1%; Score 90; DB 11; Length 588;
Best Local Similarity 34.4%; Pred. No. 6,54e-01;
Matches 11; Conservative 12; Mismatches 8; Indels 1; Gaps 1;

Db 544 tsqagvavmavavsglllllvavfvcvrrkg 575
1::: :: ::|:|:|:|: : |||:|
QY 23 TAEAAIGILTVILGVLLLI-GCWYRRRNG 53

RESULT 6
ENTRY S51663 #type complete
TITLE Lutheran blood group glycoprotein precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-May-1995 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995
S51663
S51663
AUTHORS Parsons, S.F.; Mallinson, G.; Holmes, C.H.; Houlihan, J.M.;
Simpson, K.L.; Spurr, N.K.; Warne, D.; Barclay, A.N.;
Anstee, D.J.
SUBMISSION submitted to the EMBL Data Library, December 1994
DESCRIPTION The Lutheran blood group glycoprotein, a new member of the
immunoglobulin superfamily, is widely expressed in human
tissues and is developmentally regulated in human liver.
ACCESSION S51663
STATUS preliminary
#molecule_type DNA
#residues 1-628 ##label PAR
##cross-references EMBL:X83425
SUMMARY #length 628 #molecular-weight 67374 #checksum 416

Query Match 10.1%; Score 90; DB 11; Length 628;
Best Local Similarity 34.4%; Pred. No. 6,54e-01;
Matches 11; Conservative 12; Mismatches 8; Indels 1; Gaps 1;

Db 544 tsqagvavmavavsglllllvavfvcvrrkg 575
1::: :: ::|:|:|:|: : |||:|
QY 23 TAEAAIGILTVILGVLLLI-GCWYRRRNG 53

RESULT 7
ENTRY JQ1674 #type complete
TITLE receptor protein kinase TMK1 (EC 2.7.1.1-) precursor -
Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
09-Sep-1994
JQ1674
JQ1674
AUTHORS Chang, C.; Schaller, G.E.; Patterson, S.E.; Kwok, S.F.;
Meyerowitz, E.M.; Blecker, A.B.
JOURNAL Plant Cell (1992) 4:1263-1271
TITLE The TMK1 gene from Arabidopsis codes for a protein with
structural and biochemical characteristics of a receptor
protein kinase.
ACCESSION JQ1674
#molecule_type DNA
#residues 1-942 ##label CHA
##cross-references GB:I00670

```

KEYWORDS	
glycoprotein; phosphotransferase; transmembrane protein	
FEATURE	
1-22	#domain signal sequence #status predicted #label SIG\
23-942	#product receptor protein kinase TMK1 #status predicted #label MAT\
480-503	#domain transmembrane #label TMM\
717-722	#region serine/threonine protein kinase homology\
755-763	#region serine/threonine protein kinase homology\
86,99,158,164,171, 363,533,587	#binding site carbohydrate (Asn) (covalent) #status predicted
SUMMARY	
Query Match	#length 942 #molecular-weight 102387 #checksum 2851
Best Local Similarity	9.8%; Score 88; DB 9; Length 942;
Matches	9; Conservative 10; Mismatches 4; Indels 1; Gaps 1;
Ddb	
491 gllsf1l-igllvfwykykrkf 513	
I::: : : :::	
QY 31 GLTVILGVLLIGWYCRRNCY 54	
RESULT 8	
ENTRY	S41089 #type complete
TITLE	tyrosinase-related-protein-2 - human
ORGANISM	#formal name Homo sapiens #common name man
DATE	25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
ACCESSIONS	S41089
REFERENCE	S41089
authors	Bouchard, B.; del Marmol, V.; Jackson, I.J.; Cherif, D.; Dubertret, L.
journals	Eur. J. Biochem. (1994) 219:127-134
title	Molecular characterization of a human tyrosinase-related-protein-2 cDNA. Patterns of expression in melanocytic cells.
accession	S41089
status	preliminary
molecule_type	mRNA
residues	1-519 #label BOU
SUMMARY	
Query Match	#length 519 #molecular-weight 59145 #checksum 989
Best Local Similarity	9.8%; Score 87; DB 12; Length 519;
Matches	12; Conservative 11; Mismatches 10; Indels 1; Gaps 1;
Ddb	
477 mgtlvavglfvllaflqyrllrkgytplmethl 510	
: ::::: : :::	
QY 30 IGLVTLGVLLIGWYCRR-RNGVRALMDKSL 62	
RESULT 9	
ENTRY	S43510 #type complete
TITLE	tyrosinase-related protein 2 - human
ORGANISM	#formal name Homo sapiens #common name man
DATE	13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 03-Mar-1995
ACCESSIONS	S43510
REFERENCE	S43510
authors	Yokoyama, K.; Suzuki, H.; Yasumoto, K.; Tomita, Y.; Shibahara, S.
journals	Biochim. Biophys. Acta (1994) 1217:317-321
title	Molecular cloning and functional analysis of a cDNA coding for human DOPACHrome tautomerase/tyrosinase-related

	protein-2.	
#accession	S43510	
##status	preliminary	
##molecule_type	mRNA	
##residues	1-519 ##label YOK	
##cross-references	EMBL:D17547	
SUMMARY	length 519 #molecular-weight 59145 #checksum 989	
Query Match	9.8%; Score 87; DB 12; Length 519;	
Best Local Similarity	35.3%; Pred. No. 1.58e+00;	
Matches	12; Conservative 11; Mismatches 10; Indels 1; Gaps 1;	
Dn	477 mgtlvalvgllfllaflgyrrlrkkytplmethl 510 :: :: :: :: : ::	
Qy	30 IGILTVLGVLLIGWCYRR-RNGYRALMDKSL 62	
RESULT	10	
ENTRY	D41214 #type complete	
TITLE	protein-tyrosine-phosphatase (EC 3.1.3.48) 10D, receptor-like, short splice form precursor - fruit fly (Drosophila melanogaster)	
ORGANISM	#formal_name Drosophila melanogaster	
DATE	28-May-1992 #sequence_revision 28-May-1992 #text_change 25-Aug-1995	
ACCESSIONS	D41214	
REFERENCE	A41214	
#authors	Yang, X.; Seow, K.T.; Bahri, S.M.; Oon, S.H.; Chia, W. Cell (1991) 67:661-673	
#title	Two Drosophila receptor-like tyrosine phosphatase genes are expressed in a subset of developing axons and pioneer neurons in the embryonic CNS.	
#cross-references	MUID:92034988	
#accession	C41214	
##status	preliminary	
##molecule_type	mRNA	
##residues	1-1630 ##label YAN	
##cross-references	GB:M80465	
CLASSIFICATION	#superfamily protein-tyrosine-phosphatase homology phosphoric monoester hydrolase	
KEYWORDS		
FEATURE	1295-1515 #domain protein-tyrosine-phosphatase homology #label PTP1A #active_site Cys (phosphocysteine intermediate) #status predicted 1473 #binding_site substrate phosphate (Arg) #status predicted SUMMARY length 1630 #molecular-weight 184861 #checksum 8804	
Query Match	9.8%; Score 87; DB 10; Length 1630;	
Best Local Similarity	37.0%; Pred. No. 1.58e+00;	
Matches	17; Conservative 7; Mismatches 19; Indels 3; Gaps 3;	
Dn	1184 ysfpqtd-qdntslvaivtp-ltil-vllvtlfykrrrmr 1226 :: : : : :: :: : :: :	
Qy	10 YGYPKKGHSYTAEAGIGILTIVLGVLGICWYCRRNGR 55	
RESULT	12	
ENTRY	S19243 #type complete	
TITLE	tyrosinase-related protein TRP-2 - mouse	
ORGANISM	#formal_name Mus musculus #common_name house mouse	
DATE	13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 12-Apr-1995	
ACCESSIONS	S19243	
REFERENCE	S19243	
#authors	Jackson, I.J.; Chambers, D.M.; Tsukamoto, K.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Hearing, V.	
#journal	EMBO J. (1992) 11:527-535	
#title	A second tyrosinase-related protein, TRP-2, maps to and is mutated at the mouse slaty locus.	
#cross-references	MUID:92164640	
#accession	S19243	
##status	preliminary	
##molecule_type	mRNA	
##residues	1-517 ##label JAC	
##cross-references	EMBL:X63349	
CLASSIFICATION	#superfamily monophenol monooxygenase	
SUMMARY	length 517 #molecular-weight 58569 #checksum 6717	
Query Match	9.7%; Score 86; DB 4; Length 517;	
Best Local Similarity	41.2%; Pred. No. 2.11e+00;	
Matches	14; Conservative 9; Mismatches 10; Indels 1; Gaps 1;	
Dn	475 igilgafvlllgllaflyqrrlrkgyaplmetgl 508 : : :: : :: :: : :: :	
Qy	10 YGYPKKGHSYTAEAGIGILTIVLGVLGICWYCRRNGR 55	
RESULT	11	
ENTRY	C41214 #type complete	
TITLE	protein-tyrosine-phosphatase (EC 3.1.3.48) 10D, receptor-like, long splice form precursor - fruit fly (Drosophila melanogaster)	

GENETICS

```

SUMMARY
#length 918 #molecular-weight 99530 #checksum 3552
#status predicted
Query Match 9.6%; Score 85; DB 1; Length 918;

```

Aug 27 07:56

US-08-231-565A-2.tpr

11

Best Local Similarity 34.1%; Pred. No. 2.82e+00;
Matches 15; Conservative 11; Mismatches 14; Indels 4; Gaps 3;

Db 282 gqgh-ftevlngigillvlviatlillvtacfy-rtngivril 323

|:|:| ::|:| :|:| :|:| :|:| :|:|

Qy 16 GHGHSYTTAEAAAGIGILTIVILGVLLI--GCWYCRRRNGYRAL 57

Search completed: Tue Aug 27 08:08:31 1996
Job time : 49 secs.

WAVELENGTH

(TM)

Release 2.1d John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 08:07:11 1996; MacPar time 5.49 Seconds
373.873 Million cell updates/sec

Tabular output not generated.

Title: >US-08-231-565A-2
Description: (1-118) from US08231565A.pep
Perfect Score: 889
Sequence: 1 MPREDARHLYGYPKKGHGS.....NAPPAYEKLASQPPYPSP 118

Scoring table: PAM 150
Gap 11

Searched: 49340 seqs, 17385503 residues

Post-processing: Minimum Match 0\$
Listing first 45 summaries

Database: swiss-prot32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9

Statistics: Mean 40.794; Variance 65.678; scale 0.621

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	99	11.1	344	2	CD2 MOUSE	T-CELL SURFACE ANTIGE
2	88	9.9	942	8	TMK1 ARATH	POTATIVE RECEPTOR PRO
3	87	9.8	519	8	TYR2 HUMAN	DOPACHROME TAUTOMERAS
4	87	9.8	1630	6	PTP1 DROME	PROTEIN-TYROSINE PHOS
5	86	9.7	517	8	TYR2 MOUSE	DOPACHROME TAUTOMERAS
6	85	9.6	320	9	YB9K YEAST	HYPOTHETICAL 36.0 KD
7	85	9.6	918	1	ATH1 YEAST	PLASMA MEMBRANE ATPAS
8	84	9.4	327	2	CD1A HUMAN	T-CELL SURFACE GLYCOP
9	83	9.3	306	2	CD80 MOUSE	T LYMPHOCYTE ACTIVATI
10	82	9.2	477	4	INGR MOUSE	INTERFERON-GAMMA RECE
11	82	9.2	897	2	CYRB HUMAN	CYTOKINE RECEPTOR COM
12	82	9.2	1010	9	YB9K YEAST	HYPOTHETICAL 111.7 KD
13	81	9.1	511	3	DOPR DROME	DOPAMINE RECEPTOR PRE

14	81	9.1	553	5	MCP2 SALTY	METHYL-ACCEPTING CHEM	1.86e+00
15	81	9.1	623	6	PTR2 CANAL	PEPTIDE TRANSPORTER P	1.86e+00
16	81	9.1	855	3	ENV FIVU2	ENVELOPE POLYPROTEIN	1.86e+00
17	81	9.1	856	3	ENV FIVP5	ENVELOPE POLYPROTEIN	1.86e+00
18	81	9.1	1030	9	VPP1 CAEEL	POTATIVE CLATHRIN-COA	1.86e+00
19	80	9.0	307	9	YAC2 SCHPO	HYPOTHETICAL 33.9 KD	2.58e+00
20	79	8.9	128	5	MF18 MAIZE	MFS18 PROTEIN PRECURS	3.57e+00
21	79	8.9	248	5	MYP0 HUMAN	MYELIN P0 PROTEIN PRE	3.57e+00
22	79	8.9	546	8	VGLF RINDL	FUSION GLYCOPROTEIN P	3.57e+00
23	79	8.9	846	4	ITBX DROME	INTEGRIN BETA-SUBUNIT	3.57e+00
24	79	8.9	854	3	ENV FIVD5	ENVELOPE POLYPROTEIN	3.57e+00
25	79	8.9	1184	1	BIMC EMEVI	KINESIN-LIKE PROTEIN	3.57e+00
26	78	8.8	98	3	ELIB_PHYDR	BETA-ELICITIN DRE-BET	4.93e+00
27	78	8.8	333	2	CD1B HUMAN	T-CELL SURFACE GLYCOP	4.93e+00
28	78	8.8	538	8	VGLF TRTV	FUSION GLYCOPROTEIN P	4.93e+00
29	78	8.8	700	5	NCD DROME	CLARET SEGREGATIONAL	4.93e+00
30	78	8.8	760	8	VACB SHIFL	VACB PROTEIN.	4.93e+00
31	77	8.7	98	3	ELIB_PHYCR	BETA-ELICITIN CRYPROG	6.78e+00
32	77	8.7	120	2	CYB5_YEAST	CYTOCHROME B5.	6.78e+00
33	77	8.7	248	5	MYP0 RAT	MYELIN P0 PROTEIN PRE	6.78e+00
34	77	8.7	248	5	MYP0 MOUSE	MYELIN P0 PROTEIN PRE	6.78e+00
35	77	8.7	371	1	ALF TRYBB	FRUCTOSE-BISPHOSPHATE	6.78e+00
36	77	8.7	455	4	HNF4 RAT	HEPATOCYTE NUCLEAR FA	6.78e+00
37	77	8.7	465	4	HNF4 HUMAN	HEPATOCYTE NUCLEAR FA	6.78e+00
38	77	8.7	1148	8	VGLM_PUUMH	M POLYPROTEIN PRECURS	6.78e+00
39	77	8.7	1148	8	VGLM_PUUMK	M POLYPROTEIN PRECURS	6.78e+00
40	77	8.7	1148	8	VGLM_PUUMP	M POLYPROTEIN PRECURS	6.78e+00
41	77	8.7	1329	3	FTSK ECOLI	CELL DIVISION PROTEIN	6.78e+00
42	76	8.5	304	2	CYPK_STROM	GRANATINIC POLYKETIDE	9.29e+00
43	76	8.5	463	2	CD4 CANFA	T-CELL SURFACE GLYCOP	9.29e+00
44	76	8.5	813	8	VACB ECOLI	VACB PROTEIN.	9.29e+00
45	76	8.5	920	1	ATHP_ZYGRO	PLASMA MEMBRANE ATPAS	9.29e+00

ALIGNMENTS

RESULT	1	CD2_MOUSE	STANDARD;	PRT;	344 AA.
AC	P08920;				
DT	01-NOV-1988	(REL. 09, CREATED)			
DT	01-NOV-1988	(REL. 09, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)			
DE	T-CELL SURFACE ANTIGEN CD2 PRECURSOR (T-CELL SURFACE ANTIGEN				
DE	T11/LEU-5) (IFA-2) (IFA-3 RECEPTOR).				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-B10.A;				
RX	MEDLINE; 87271635.				
RA	SCWELL W.A., BROWN M.H., FINK P.J., KOZAK C.A., CRUMPTON M.J.;				
RL	EUR. J. IMMUNOL. 17:1015-1020(1987).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 88004738.				
RA	CLAYTON L.K., SAYRE P.H., NOVOTNY J., REINHERZ E.L.;				
RL	EUR. J. IMMUNOL. 17:1367-1370(1987).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 88144486.				
RA	DIAMOND D.J., CLAYTON L.K., SAYRE P.H., REINHERZ E.L.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 85:1615-1619(1988).				
CC	-!- FUNCTION: CD2 INTERACTS WITH LYMPHOCYTE FUNCTION-ASSOCIATED				

RA YOKOYAMA K., SUZUKI H., YASUMOTO K.I., TOMITA Y., SHIBAHARA S.;
 RL BIOCHIM. BIOPHYS. ACTA 1217:317-321(1994).
 CC -/- CATALYTIC ACTIVITY: DOPACHROME = 5, 6-DIHYDROXYINDOLE-2-CARBOXYLATE.
 CC -/- COFACTOR: CONTAINS TWO ZINC ATOMS (BY SIMILARITY).
 CC -/- PATHWAY: MELANIN BIOSYNTHESIS.
 CC -/- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MELANOSOMAL.
 CC -/- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
 DR EMBL; D17547; D17547.
 DR PIR; S43510; S43510.
 DR MIM; 191275; 11TH EDITION.
 DR PROSITE; PS00497; TYROSINASE 1.
 DR PROSITE; PS00498; TYROSINASE 2.
 RW ISOMERASE; ZINC; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;
 KW MELANIN BIOSYNTHESIS.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 519 DOPACHROME TAUTOMERASE.
 FT DOMAIN 24 472 LUMENAL, MELANOSOME (POTENTIAL).
 FT TRANSMEM 473 493 POTENTIAL.
 FT DOMAIN 494 519 CYTOPLASMIC (POTENTIAL).
 FT METAL 189 189 ZINC A (BY SIMILARITY).
 FT METAL 211 211 ZINC A (BY SIMILARITY).
 FT METAL 220 220 ZINC A (BY SIMILARITY).
 FT METAL 369 369 ZINC B (BY SIMILARITY).
 FT METAL 373 373 ZINC B (BY SIMILARITY).
 FT METAL 396 396 ZINC B (BY SIMILARITY).
 FT METAL 170 170 POTENTIAL.
 FT CARBOHYD 178 178 POTENTIAL.
 FT CARBOHYD 237 237 POTENTIAL.
 FT CARBOHYD 300 300 POTENTIAL.
 FT CARBOHYD 342 342 POTENTIAL.
 FT CARBOHYD 377 377 POTENTIAL.
 SQ SEQUENCE 519 AA; 59145 MW; 1422986 CN;

Query Match 9.8%; Score 87; DB 8; Length 519;
 Best Local Similarity 35.3%; Pred. No. 2.41e-01;
 Matches 12; Conservative 11; Mismatches 10; Indels 1; Gaps 1;

Db 477 mgltvalvqlfvlafqyrrlrkgytplmethl 510
 :| |:::|::| | |::| |:
 Qy 30 IGLTVILVILLICWYCR-RNGYRALMDKSL 62

RESULT 4
 ID PTP1 DROME STANDARD; PRT; 1630 AA.
 AC P35992;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE PROTEIN-TYROSINE PHOSPHATASE 10D PRECURSOR (EC 3.1.3.48) (RECEPTOR-LINKED PROTEIN-TYROSINE PHOSPHATASE 10D).
 GN PTP10D.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 92034989.
 RA TAN S.-S., TSOUFAS P., ZINN K.;
 RL CELL 67:675-685(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 92034988.

RA YANG X., SEOW K.T., BAHRI S.M., OON S.H., CHIA W.;
 RL CELL 67:661-673(1991).
 CC -/- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
 CC -/- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -/- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED IN A SUBSET OF AXONS AND PIONEER NEURONS IN THE EMBRYO.
 CC -/- SIMILARITY: CONTAINS ONE PROTEIN-TYROSINE PHOSPHATASE DOMAIN AND 12 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -/- ALTERNATIVE PRODUCTS: TWO ISOFORMS THAT DIFFER IN THEIR C-TERMINAL TAILS ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC EMBL; M80538; M80538.
 DR EMBL; M80465; M80465.
 DR PIR; C41214; C41214.
 DR PIR; D41214; D41214.
 DR PIR; A41215; A41215.
 DR HSP; P02751; 1TF.
 DR FLYBASE; FBGN0004370; PTP10D.
 DR PROSITE; PS00383; TYR. PHOSPHATASE.
 KW TRANSMEMBRANE; HYDROLASE; DUPLICATION; SIGNAL; ALTERNATIVE SPLICING;
 KW REPEAT.
 FT SIGNAL 1 34 POTENTIAL.
 FT CHAIN 35 1630 PROTEIN-TYROSINE PHOSPHATASE 10D.
 FT DOMAIN 35 1196 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1197 1218 POTENTIAL.
 FT DOMAIN 1219 1630 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 43 119 FIBRONECTIN TYPE-III.
 FT DOMAIN 120 214 FIBRONECTIN TYPE-III.
 FT DOMAIN 215 308 FIBRONECTIN TYPE-III.
 FT DOMAIN 309 402 FIBRONECTIN TYPE-III.
 FT DOMAIN 403 493 FIBRONECTIN TYPE-III.
 FT DOMAIN 494 580 FIBRONECTIN TYPE-III.
 FT DOMAIN 581 669 FIBRONECTIN TYPE-III.
 FT DOMAIN 670 766 FIBRONECTIN TYPE-III.
 FT DOMAIN 767 861 FIBRONECTIN TYPE-III.
 FT DOMAIN 862 955 FIBRONECTIN TYPE-III.
 FT DOMAIN 956 1048 FIBRONECTIN TYPE-III.
 FT DOMAIN 1049 1189 FIBRONECTIN TYPE-III.
 FT DOMAIN 1290 1532 PROTEIN-TYROSINE PHOSPHATASE.
 FT CARBOHYD 75 75 POTENTIAL.
 FT CARBOHYD 106 106 POTENTIAL.
 FT CARBOHYD 128 128 POTENTIAL.
 FT CARBOHYD 169 169 POTENTIAL.
 FT CARBOHYD 212 212 POTENTIAL.
 FT CARBOHYD 229 229 POTENTIAL.
 FT CARBOHYD 259 259 POTENTIAL.
 FT CARBOHYD 289 289 POTENTIAL.
 FT CARBOHYD 317 317 POTENTIAL.
 FT CARBOHYD 471 471 POTENTIAL.
 FT CARBOHYD 486 486 POTENTIAL.
 FT CARBOHYD 512 512 POTENTIAL.
 FT CARBOHYD 533 533 POTENTIAL.
 FT CARBOHYD 588 588 POTENTIAL.
 FT CARBOHYD 668 668 POTENTIAL.
 FT CARBOHYD 687 687 POTENTIAL.
 FT CARBOHYD 719 719 POTENTIAL.
 FT CARBOHYD 723 723 POTENTIAL.
 FT CARBOHYD 823 823 POTENTIAL.
 FT CARBOHYD 841 841 POTENTIAL.
 FT CARBOHYD 874 874 POTENTIAL.
 FT CARBOHYD 908 908 POTENTIAL.
 FT CARBOHYD 925 925 POTENTIAL.
 FT CARBOHYD 1001 1001 POTENTIAL.
 FT CARBOHYD 1104 1104 POTENTIAL.

Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Db 287 fgvlslvlpfymgscwyrkw 310

:||: || | : ||| ||| :

Qy 30 IGLTWILGV-LLLIGC-WYCRRR 51

RESULT 7

ID ATH1 YEAST STANDARD; PRT; 918 AA.

AC P05030;

DT 13-AUG-1987 (REL. 05, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE PLASMA MEMBRANE ATPASE 1 (EC 3.6.1.35) (PROTON PUMP).

GN PMA1.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 86146844.

RA SERRANO R., KIELLAND-BRANDT M.C., FINK G.R.;

RL NATURE 319:689-693(1986).

CC -/- FUNCTION: THE PLASMA MEMBRANE ATPASE OF PLANTS AND FUNGI IS A

CC HYDROGEN ION PUMP. THE PROTON GRADIENT IT GENERATES DRIVES THE

CC ACTIVE TRANSPORT OF NUTRIENTS BY H⁺-SYMPORT. THE RESULTING

CC EXTERNAL ACIDIFICATION AND/OR INTERNAL ALKINIZATION MAY MEDIATE

CC GROWTH RESPONSES.

CC -/- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.

CC -/- THERE ARE TWO PLASMA MEMBRANE ATPASES IN YEAST. THIS IS THE

CC MAJOR ISOFORM.

CC -/- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -/- PTM: PHOSPHORYLATED ON MULTIPLE SER AND THR RESIDUES.

CC -/- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY

CC (EI-E2 ATPASES).

DR EMBL; X03534; X03534.

DR PIR; A25823; PXBYIP.

DR LISTA; SC00832; PMA1.

DR SGD; I0001449; PMA1.

DR PROSITE; PS00154; ATPASE_E1_E2.

KW HYDROLASE; HYDROGEN ION TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;

KW ATP-BINDING; MULTIGENE FAMILY; GLYCOPROTEIN.

FT DOMAIN 1 115

FT TRANSMEM 116 136

FT DOMAIN 137 140

FT TRANSMEM 141 160

FT DOMAIN 161 291

FT TRANSMEM 292 313

FT DOMAIN 314 325

FT TRANSMEM 325 347

FT DOMAIN 348 719

FT TRANSMEM 720 738

FT DOMAIN 739 754

FT TRANSMEM 755 774

FT DOMAIN 775 824

FT TRANSMEM 825 845

FT DOMAIN 846 857

FT TRANSMEM 858 874

FT DOMAIN 875 918

FT MOD_RES 378 378

FT BINDING 474 474

FT CARBOHYD 848 848

FT DOMAIN 15 25

FT DOMAIN 31 78

SQ SEQUENCE 918 AA; 99619 MW; 4420147 CN;

Query Match 9.6%; Score 85; DB 1; Length 918;
Best Local Similarity 34.1%; Pred. No. 4.82e-01;
Matches 15; Conservative 11; Mismatches 14; Indels 4; Gaps 3;

Db 282 gqgh-ftevlmgigillvltatlllwtacfy-rtnqivril 323

:||: || : ||| :||| :||| :

Qy 16 GHGSYTTAEAGIGILTIVILGVLLLI--GWCYRRRNGYRAL 57

RESULT 8

ID CDIA HUMAN STANDARD; PRT; 327 AA.

AC P06126;

DT 01-JAN-1988 (REL. 06, CREATED)

DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)

DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)

DE T-CELL SURFACE GLYCOPROTEIN CDIA PRECURSOR (CDIA ANTIGEN) (T-CELL

DE SURFACE ANTIGEN T6/LEU-6) (HTAI THYMOCYTE ANTIGEN).

GN CDIA.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 88097453.

RA MARTIN L.H., CALABI F., LEFEBVRE F.-A., BILSLAND C.A.G., MILSTEIN C.;

RL PROC. NATL. ACAD. SCI. U.S.A. 84:9189-9193(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 89341413.

RA ARUFFO A., SEED B.;

RL J. IMMUNOL. 143:1723-1730(1989).

RN [3]

RP SEQUENCE OF 99-327 FROM N.A.

RC TISSUE=T-CELL;

RX MEDLINE; 87014824.

RA CALABI F., MILSTEIN C.;

RL NATURE 323:540-543(1986).

CC -/- FUNCTION: NOT KNOWN.

CC -/- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -/- SUBUNIT: ASSOCIATES NON-COVALENTLY WITH BETA-2-MICROGLOBULIN.

CC -/- TISSUE SPECIFICITY: EXPRESSED ON CORTICAL THYMOCYTES, ON CERTAIN

CC T-CELL LEUKEMIAS, AND IN VARIOUS OTHER TISSUES.

CC -/- SIMILARITY: TO OTHER CD1 ANTIGENS, AND TO MHC CLASS I ANTIGENS.

CC -/- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

DR EMBL; M28825; M28825.

DR EMBL; M22080; M22080.

DR EMBL; M22163; M22163.

DR EMBL; M22164; M22164.

DR EMBL; M22165; M22165.

DR EMBL; M22166; M22166.

DR EMBL; M22167; M22167.

DR EMBL; X04450; X04450.

DR PIR; A02242; HLHUCD.

DR PIR; A39957; A39957.

DR PIR; A45801; A45801.

DR HSP; P01842; IDLH.

DR MM; 188370; 11TH EDITION.

KW GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE; IMMUNOGLOBULIN FOLD; T-CELL;

KW MULTIGENE FAMILY.

FT SIGNAL 1 16

FT CHAIN 17 327

FT DOMAIN 17 108

FT DOMAIN 109 201

CDIA ANTIGEN.

EXTRACELLULAR ALPHA-1.

EXTRACELLULAR ALPHA-2.

Aug 27 07:55

US-08-231-565A-2.rsp

11

FT DOMAIN 202 294 EXTRACELLULAR ALPHA-3.
 FT TRANSMEM 300 321 POTENTIAL.
 FT DOMAIN 322 327 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 119 183 BY SIMILARITY.
 FT CARBOHYD 37 37 POTENTIAL.
 FT CARBOHYD 60 60 POTENTIAL.
 FT CARBOHYD 74 74 POTENTIAL.
 FT CARBOHYD 145 145 POTENTIAL.
 FT CONFLICT 30 30 I -> T (IN REF. 1).
 FT CONFLICT 68 68 W -> C (IN REF. 1).
 SQ SEQUENCE 327 AA; 37172 MW; 559220 CN;

Query Match 9.4%; Score 84; DB 2; Length 327;
 Best Local Similarity 48.1%; Pred. No. 6.78e-01;
 Matches 13; Conservative 7; Mismatches 4; Indels 3; Gaps 3;

Db 299 svgfllavlpvlllglalwf-kr 324
 QY : |: ||: ||: ||: ||: ||: |: |: |:
 27 AAGIGILTVILGVLLIG-C-WYCR 51

RESULT 9

ID CD80 MOUSE STANDARD; PRT; 306 AA.
 AC Q00609;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE T LYMPHOCYTE ACTIVATION ANTIGEN CD80 PRECURSOR (ACTIVATION B7-1
 DE ANTIGEN) (B7).
 GN CD80 OR B7.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=B-CELL;
 RX MEDLINE; 91341422.
 RA GRAY G.S., FREEMAN G.J., GIMMI C.D., LOWARD D.B., ZHOU L.J.,
 RA WHITE M., FINGEROTH J.D., GRIBBEN J.G., NADLER L.M.;
 RL J. EXP. MED. 174:625-631 (1991).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=B-CELL;
 RX MEDLINE; 93307789.
 RA SELVAKUMAR A., WHITE P.C., DUPONT B.;
 RL IMMUNOGENETICS 38:292-295 (1993).
 CC -!- FUNCTION: INVOLVED IN THE COSTIMULATORY SIGNAL ESSENTIAL FOR T
 CC LYMPHOCYTES ACTIVATION. T CELL PROLIFERATION AND CYTOKINE
 CC PRODUCTION IS INDUCED BY THE BINDING OF CD28 OR CTLA-4 TO THIS
 CC RECEPTOR.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED BETWEEN 4 AND 12 HOURS POST-
 CC ACTIVATION. PROTEIN WAS DETECTED AT CELL SURFACE AT 24 HOURS AND
 CC IT'S EXPRESSION WAS MAXIMAL FROM 48 TO 72 HOURS POST-ACTIVATION.
 CC -!- TISSUE SPECIFICITY: EXPRESSED ON ACTIVATED B CELLS, GAMMA
 CC INTERFERON STIMULATED MONOCYTES AND NONCIRCULATING B-CELL
 CC MALIGNANCIES.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
 CC ONE C2-LIKE AND ONE V-LIKE DOMAIN.
 DR EMBL; X60958; X60958.
 DR EMBL; L12587; L12585.
 DR EMBL; L12586; L12586.
 DR EMBL; L12587; L12587.
 DR EMBL; L12588; L12588.

Aug 27 07:55

US-08-231-565A-2.rsp

12

DR EMBL; L12589; L12589.
 DR PIR; S17291; S17291.
 KW IMMUNOGLOBULIN FOLD; T-CELL; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE.
 FT SIGNAL 1 37
 FT CHAIN 38 306
 FT DOMAIN 22 246
 FT TRANSMEM 247 268
 FT DOMAIN 269 306
 FT DOMAIN 47 126
 FT DOMAIN 158 226
 FT DOMAIN 227 246
 FT DISULFID 54 119
 FT DISULFID 165 219
 FT CARBOHYD 93 93
 FT CARBOHYD 99 99
 FT CARBOHYD 149 149
 FT CARBOHYD 189 189
 FT CARBOHYD 210 210
 FT CARBOHYD 214 214
 SQ SEQUENCE 306 AA; 34589 MW; 505533 CN;

Query Match 9.3%; Score 83; DB 2; Length 306;
 Best Local Similarity 27.6%; Pred. No. 9.52e-01;
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 252 gagfgavtvtvvlvilk-fckhrscfr 279
 QY : ||: |: : : ||: |: |: |: |:
 27 AAGIGILTVILGVLLIGCWYCRNGYR 55

RESULT 10

ID INGR MOUSE STANDARD; PRT; 477 AA.
 AC P15261;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE INTERFERON-GAMMA RECEPTOR ALPHA CHAIN PRECURSOR.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90083245.
 RA MUNRO S., MANIATIS T.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:9248-9252 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90046824.
 RA GRAY P.W., LEONG S., FENNIE E.H., FARRAR M.A., PINGEL J.T.,
 RA FERNANDEZ-LUNA J., SCHREIBER R.D.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:8497-8501 (1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90099370.
 RA HEMMI S., PEGHINI P., METZLER M., MERLIN G., DEMBIC Z., AGUET M.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:9901-9905 (1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90036866.
 RA KUMAR C.S., MUTHUKUMARAN G., FROST L.J., NOE M., AHN Y.H.,
 RA MARIANO T.M., PESTKA S.;
 RL J. BIOL. CHEM. 264:17939-17946 (1989).
 RN [5]
 RP SEQUENCE FROM N.A.

```

RX MEDLINE; 90154099.
RA COFANO F., MOORE S.K., TANAKA S., YUHKI N., LANDOLFO S.,
RA APPELLA E.;
RL J. BIOL. CHEM. 265:4064-4071(1990).
CC -I- FUNCTION: RECEPTOR FOR INTERFERON GAMMA. TWO RECEPTORS BIND ONE
CC INTERFERON-GAMMA DIMER.
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I- PTM: PHOSPHORYLATED AT SER/THR RESIDUES.
CC -I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC TWO IG-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -I- SIMILARITY: BELONGS TO THE CLASS II CYTOKINE FAMILY OF RECEPTORS.
DR EMBL; M28995; M28995.
DR EMBL; M26711; M26711.
DR EMBL; M28233; M28233.
DR EMBL; M25764; M25764.
DR EMBL; J05265; J05265.
DR PIR; A34368; A34368.
DR PIR; A34423; A34423.
DR PIR; A34508; A34508.
DR PIR; A35468; A35468.
DR PIR; A36224; A36224.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; PHOSPHORYLATION;
KW IMMUNOGLOBULIN FOLD.
FT SIGNAL 1 22
FT CHAIN 23 477 INTERFERON-GAMMA RECEPTOR ALPHA CHAIN.
FT DOMAIN 23 253 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 254 277 POTENTIAL.
FT DOMAIN 278 477 CYTOPLASMIC (POTENTIAL).
FT DISULFID 83 91 BY SIMILARITY.
FT DISULFID 128 174 BY SIMILARITY.
FT DISULFID 203 208 BY SIMILARITY.
FT DISULFID 222 243 BY SIMILARITY.
FT CARBOHYD 61 61 POTENTIAL.
FT CARBOHYD 85 85 POTENTIAL.
FT CARBOHYD 186 186 POTENTIAL.
FT CARBOHYD 204 204 POTENTIAL.
FT CARBOHYD 211 211 POTENTIAL.
SQ SEQUENCE 477 AA; 52271 MW; 1295958 CN;

Query Match 9.2%; Score 82; DB 4; Length 477;
Best Local Similarity 36.1%; Pred. No. 1.33e+00;
Matches 13; Conservative 13; Mismatches 6; Indels 4; Gaps 4;

Db 260 vapltvftvtilvfaywy-ttknsfrksimlpksl 294
:: ||||: |::: || ::||: |:: |||
QY 30 IGLITVLGVLILLICWICVRRNGY-R-ALM-DKSL 62

RESULT 11
ID CYR6 HUMAN STANDARD; PRT; 897 AA.
AC P32927;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR.
GN CSF2RB OR IL5RB OR IL3RB.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RX MEDLINE; 9108571.
RA HAYASHIDA K., KITAMURA T., GORMAN D.M., ARAI K., YOKOTA T.,

```

```

RA MIYAJIMA A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:9655-9659(1990).
CC -I- FUNCTION: HIGH AFFINITY RECEPTOR FOR INTERLEUKIN-3, INTERLEUKIN-5
CC AND GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR.
CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. THE BETA
CC CHAIN IS COMMON TO THE IL-3, IL-5 AND GM-CSF RECEPTORS.
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -I- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
DR EMBL; M59941; M59941.
DR PIR; A39255; A39255.
DR MIM; 138981; 11TH EDITION.
DR PROSITE; PS00241; RECEPTOR_CYTOKINES_1.
DR RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 897 CYTOKINE RECEPTOR COMMON BETA CHAIN.
FT DOMAIN 17 443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 444 460 POTENTIAL.
FT DOMAIN 461 897 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 129 238 FIBRONECTIN TYPE-III.
FT DOMAIN 336 434 FIBRONECTIN TYPE-III.
FT DISULFID 35 45 BY SIMILARITY.
FT DISULFID 75 91 BY SIMILARITY.
FT CARBOHYD 58 58 POTENTIAL.
FT CARBOHYD 191 191 POTENTIAL.
FT CARBOHYD 346 346 POTENTIAL.
SQ SEQUENCE 897 AA; 97323 MW; 4443497 CN;

Query Match 9.2%; Score 82; DB 2; Length 897;
Best Local Similarity 33.3%; Pred. No. 1.33e+00;
Matches 14; Conservative 14; Mismatches 10; Indels 4; Gaps 4;

Db 528 qfgdsevepltkiedkhvcdpsqpsdttpaasdlpteqppsp 569
|| : | :::: | :| |:: || ::||:|:|
QY 78 GFDHRD-SKVSIOQE-KN-CEPVV-PNAPPAYEKLSEQSPPP 115

RESULT 12
ID YRK6 YEAST STANDARD; PRT; 1010 AA.
AC P38163;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 111.7 KD PROTEIN IN PKC1 5'REGION.
GN YBL106C OR YBL0806.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA OBERMAIER B., GASSENHUBER J., PIRAVANDI E., DOMDEY H.;
RL YEAST 11:1103-1112(1995).
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
DR EMBL; X79489; X79489.
DR EMBL; Z35867; Z35867.
DR PIR; S45389; S45389.
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT TRANSMEM 107 123 POTENTIAL.
FT TRANSMEM 363 379 POTENTIAL.
FT TRANSMEM 608 626 POTENTIAL.
FT TRANSMEM 763 779 POTENTIAL.
FT TRANSMEM 789 806 POTENTIAL.
SQ SEQUENCE 1010 AA; 111678 MW; 5433252 CN;

```

Query Match 9.1%; Score 81; DB 3; Length 511;
Best Local Similarity 36.7%; Pred. No. 1.86e+00;
Matches 11; Conservative 9; Mismatches 7; Indels 3; Gaps 3;

b
y

380 hsspyhvsdhkaavt-gvimgvfl-i-cw 406
| : | :: | :: : | :: | : | ||
17 HGHSYTTAEFAAGIGILTIVLGVLLIGCW 46

RESULT		14		
ID	MCP2 SALTY STANDARD;	PRT;	553 AA.	
AC	P02941;			
DT	21-JUL-1986 (REL. 01,	CREATED)		
DT	01-NOV-1995 (REL. 32,	LAST SEQUENCE UPDATE)		
DT	01-NOV-1995 (REL. 32,	LAST ANNOTATION UPDATE)		
DE	DE METHYL-ACCEPTING CHEMOTAXIS PROTEIN II (MCP-II) (ASPARTATE CHEMORECEPTOR PROTEIN).			
GN	TAR.			
OS	SALMONELLA TYPHIMURIUM.			
OC	PROKARYOTA; GRACILICUTES;	SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS; ENTEROBACTERIACEAE.		
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; #3197387.			
RA	RUSSO A.F., KOSHLAND D.E. JR.; SCIENCE 220:1016-1020(1983).			
RL				
RN	[2]			
RP	REVISIONS TO 236 AND INSERTION OF ALA-498. STRAIN=LT2;			
RC				
RA	STOCK A.;			
RL	THESTS (1986), UNIVERSITY OF BERKELEY, U.S.A.			
RN	[3]			

RE A-11 CRYSTALLINITY: 12-10%
 YEH J.I., BIEHMANN H.-P., PANDIT J., KOSHIAID D.E., KIM S.-H.;
 J. BIOL. CHEM. 268:9787-9792(1993).
 CC -I- FUNCTION: RECEPTOR FOR THE ATTRACTANT L-ASPARTATE AND RELATED
 CC AMINO AND DICARBOXYLIC ACIDS. TAR MEDIATES TAXIS AWAY FROM THE
 CC REPELLENTS COBALT AND NICKEL. UNLIKE E.COLI TAR IT DOES NOT
 CC MEDIATES MALTOSE TAXIS.
 CC -I- FUNCTION: CHEMOTACTIC-SIGNAL TRANSDUCERS RESPOND TO CHANGES IN THE
 CC CONCENTRATION OF ATTRACTANTS AND REPELLENTS IN THE ENVIRONMENT, AND
 CC TRANSDUCE A SIGNAL FROM THE OUTSIDE TO THE INSIDE OF THE CELL, AND
 CC FACILITATE SENSORY ADAPTATION THROUGH THE VARIATION OF THE LEVEL
 CC OF METHYLATION. ATTRACTANTS INCREASE THE LEVEL OF METHYLATION
 CC WHILE REPELLENTS DECREASE THE LEVEL OF METHYLATION, THE METHYL
 CC GROUPS ARE ADDED BY THE METHYLTRANSFERASE CHER AND REMOVED BY THE
 CC METHYLSTERASE CHEB.
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE..
 CC -I- SIMILARITY: TO THE OTHER BACTERIAL CHEMOTAXIS SENSORY TRANSDUCERS.
 DR EMBL; J01809; J01809.
 DR PIR; A03441; QREB0T.
 DR PDB; 1WAS; 15-JAN-95.
 DR PDB; 1WAT; 15-JAN-95.
 DR STYGENE; SG10387; TAR.
 DR PROSITE; P500538; CHEMOTAXIS_TRANSDUCER.
 KW CHEMOTAXIS; TRANSDUCER; TRANSMEMBRANE; METHYLATION; PERIPLASMIC;
 KW OUTER MEMBRANE; 3D-STRUCTURE.
 FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 7 33 POTENTIAL..
 FT CHAIN 34 552 TAR PROTEIN.
 FT DOMAIN 34 190 PERIPLASMIC (POTENTIAL).
 FT TRANSEM 191 215 POTENTIAL.
 FT DOMAIN 196 552 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 64 73 THE 3 ARG MAY FORM A POSITIVELY CHARGED

THE 3 ARG MAY FORM A POSITIVELY CHARGED CLOUSTERIC (FOUNTAIN).

Aug 27 07:55

US-08-231-565A-2.jsp

17

POCKET, WHICH BINDS THE ALPHA-CARBOXYL

GROUP OF THE ATTRACTANT AA.

DEAMINATION AND METHYLATION.

METHYLATION.

DEAMINATION AND METHYLATION.

METHYLATION.

SEQUENCE 553 AA; 59511 MW; 1383542 CN;

Query Match 9.1%; Score 81; DB 5; Length 553;

Best Local Similarity 40.9%; Pred. No. 1.86e+00;

Matches 9; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

Db 194 lgvlavvl-vlllmvfwfgrh 214

:|:|:| | | | | | | | | |

QY 30 IGLTVILGVLLIGCWYCRRR 51

RESULT 15

ID PTR2 CANAL STANDARD; PRT; 623 AA.

AC P46030;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE PEPTIDE TRANSPORTER PTR2.

GN PTR2.

OS CANDIDA ALBICANS (YEAST).

OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).

RP SEQUENCE FROM N.A.

RC STRAIN-CBS 562 / ATCC 18804;

RA BASRAI M.A., LUBKOWITZ M.A., PERRY J.R., MILLER D., KRAINER E.,

RA NAIDER F.R., BECKER J.M.;

RL MICROBIOLOGY 141:1147-1156(1995).

CC -1- FUNCTION: UPTAKE OF SMALL PEPTIDE.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.

DR EMBL; U09781; U09781.

DR PROSITE; PS01022; PTR2_1.

DR PROSITE; PS01023; PTR2_2.

KW PEPTIDE TRANSPORT; TRANSPORT; TRANSMEMBRANE.

FT TRANSMEM 134 154 POTENTIAL.

FT TRANSMEM 163 183 POTENTIAL.

FT TRANSMEM 191 211 POTENTIAL.

FT TRANSMEM 250 270 POTENTIAL.

FT TRANSMEM 277 297 POTENTIAL.

FT TRANSMEM 385 405 POTENTIAL.

FT TRANSMEM 418 438 POTENTIAL.

FT TRANSMEM 448 468 POTENTIAL.

FT TRANSMEM 499 519 POTENTIAL.

FT TRANSMEM 529 549 POTENTIAL.

FT TRANSMEM 557 577 POTENTIAL.

SQ SEQUENCE 623 AA; 69941 MW; 2021848 CN;

Query Match

Best Local Similarity 39.1%; Score 81; DB 6; Length 623;

Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 256 lainigafqlqiatycoerrvgfw 278

:|:|:| | | | | | | | | |

QY 33 LTVILGVLLIGCWYCRRCYR 55

Search completed: Tue Aug 27 08:07:23 1996

Job time : 12 secs.

(TM)

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
	1	1392	98.0	1559	17	T02714	MART-1 melanoma anti	1.35e-108
	2	1372	96.6	354	12	Q76370	Tumour rejection anti	8.83e-107
	3	1332	93.8	760	17	T11550	Tumour rejection anti	3.77e-103
	4	1443	31.2	13585	17	T11549	Tumour rejection anti	2.85e-24
c	5	159	11.2	416	4	N10053	Generic coding sequen	9.55e-02
c	6	157	11.1	231	2	N60354	Desulphatohirudin (I)	1.31e-01
	7	138	9.7	1467	3	Q20927	Leptospira hardjo-bov	2.47e+00
	8	132	9.3	3556	11	O63952	VEF gene.	6.04e+00

PT immunise animal against melanoma.
PS Claim 2; Page 115-116; 184pp; English.
CC The nucleic acid encodes a melanoma antigen (MART-1) which is
CC recognized by T-lymphocytes. It is used for recombinant protein
CC production, preferably using a baculo virus vector for expression in
CC insect cell cultures. MART-1 protein is a source of immunogenic
CC peptides (see R84196 for peptide M9-2) which are optionally modified
CC (see R84783-R84800) and used in medicaments for the treatment or
CC prevention (by immunization) of melanoma. Antibodies against MART-1
CC and its immunogenic peptides may be used in the detection and
CC isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 1559 BP; 470 A; 330 C; 324 G; 435 T;

Query Match	98.0%;	Score 1392;	DB 17;	Length 1559;
Best Local Similarity	58.6%;	Pred. No. 1.35e-108;		
Matches	207;	Conservative 81;	Mismatches 65;	Indels 0;
		Gaps		

	Ddb	54	atgccagaagaaagtctcacttcatctatcggtaccaccaagaaggggcacgccaacctt113 : : : : : : : : : : : :
	Oy	1	ATGCNNMGARGAYGCGNCAYYTHTAYGNTTCYCNNAARAARGCNCAYGCGNCAWSN60
	Ddb	114	tacaccacgcgtgaagagccgcctgggatcgcatcctgacatcgatctggagtgctta173 : : : : : : : : : : : :
	Oy	61	TAYACNACNGCARGAGCGCGGNATHGGNATHYNACNGTNATHYTGNGGTNYTN120
	Ddb	174	ctgctcatcgctgttgattgtagaagcaaatggatacagacgcttgatgataaa.233 : : : : : : : : : : : : :
	Oy	121	YTYNYNATHGGCTGYGTGTATGYMGNMNQHAAYCGNTAYMGCNGCNTYTNATGGAUAR180
	Ddb	234	agtcttcattctggcactcaatgtgctttacaagaagatgccacaagaagggtttgtat293 : : : : : : : : : : : :
	Oy	181	WSNYNCAYTTGGWAGNCARTGTGCNTACNMGMNWGTGYCCNCARGAGNTTYGAY240
	Ddb	294	catcggaagacagaagtgtctctcaagaaaaactgtgaacctctggttcccacatgct353 : : : : : : : : : : : :
	Oy	241	CAYMNGAYWSAARGTWNSNYTNCARGAARAAYTGYGARCCNGTGTGCCNAAAYCCN300
	Ddb	354	ccacctgttatgaaaactctctgcagagaacgtcacaccaccttatcaac406 : : : : : : : : : : : :
	Oy	301	CNCNCNTATYGAAAARYTNWGNGCARCWSCNCNCNCNTATYASNOCC353

[illegible]

PI	Boon-Falloux T, Brichard V, De Plaen E, Traversari C;
PI	Van Pel A, Wolfel T;
DR	WPI; 94-316544/39.
DR	P-PSDB; R63158.
PT	Nucleic acid coding for a tumour rejection antigen precursor - is
PT	used for developing prods. for diagnosis or treatment of expression
PT	related disorders, partic. melanoma
PS	Claim 5; Page 14; 26pp; English.
CC	This sequence encodes the tumour rejection antigen precursor which is
CC	processed to a tumour rejection antigen presented by HLA-A2 molecules.
CC	The tumour rejection antigen is not related to tyrosinase. This
CC	sequence was isolated from the melanoma cell line, LB-39-MEL. The
CC	tumour rejection antigen may be used for diagnosis or in vaccines or
CC	for therapy of disorders characterised by the expression of the tumour
CC	rejection antigen precursor, particularly melanoma.
SQ	Sequence 354 BP; 99 A; 86 C; 86 G; 83 T;

Query Match	96.6%;	Score 1372;	DB 12;	Length 354;
Best Local Similarity	58.1%;	Pred. No. 8,836-107;		
Matches	205;	Conservative 81;	Mismatches 67;	Indels 0;
Db	1	atcgagaagaagatgctcattctcatctatggttacccaagaaggagacggccaactt	60	
Qy	1	ATGCCNMGNCARGACVGCNCATYTTATHYAGNTAYTCNNAARAAGNCAYGNCATWSN	60	
Db	61	tacacacggctgaagagcgctggatcgagctctgcagctgatcctggagcttta	120	
Qy	61	TAYACNACGNCARGARGCNGCGGNATHGGNATHYTNACNGTNTATHTNGGNGTNTN	120	
Db	121	ctgctcatcggctgttgattatgaagaagcaaatggatcacagagccttgatggataaa	180	
Qy	121	YTYNTNATHGNTGYTGCTATGYCMNGMNGNAAYCGNTAYMGNGTNTATGGAAYAR	180	
Db	181	agttcttcattgtggcactcaatgctcttaagaagaagatgccacaagaaggttgat	240	
Qy	181	WSNYNCAYGTNGNACNCARTGYGNTYNACNMGWNGTGYCCNCARGAGGNTTYGAY	240	
Db	241	catcgggacagcaaatgctcttcaadgaagaaactgtgaacctgtgggtcccaatgct	300	
Qy	241	CATYMGNGAYWSNNAARGTNWSYNTNCARGAARAATYGTGARCCNGTNGTCCNNAYGON	300	
Db	301	caacctgcttatgaagaactctctgcagaacagtccaccacacttattcacc	353	
Qy	301	CNCCNGNTATYGAARAATYTNWNGNARGCARISNCCNCCNCCNTATYNSCC	353	

RESULT	3
ID	T11550 standard; cDNA; 760 BP.
AC	T11550;
DE	18-APR-1996 (first entry)
DT	Tumour rejection antigen precursor coding sequence.
DD	Tumour rejection antigen precursor; TRAP; TRA; melanoma; cancer;
KW	tumour; treatment; detection; vaccine; HLA-A2; adoptive transfer;
KW	T cell; T lymphocyte; human leukocyte antigen; as.
KW	Homo sapiens.
OS	
PN	M09601557-Al.
PD	25-JAN-1996.
PP	27-JUN-1995; U08153.
PP	08-JUL-1994; US-272351.
PR	10-JAN-1995; US-370319.
PA	(LUDW-) LUDWIG INST CANCER RES.
PA	Boon-fallier T, Brichard V, Coulie P, De PLAEN E;
PPI	Traversari C, Van PEL A, Wolfel T;
PI	Immunogenetics 1996; 48(1):1-10.
DR	WFJ; 96-097390/10.

FT	/note=	"encodes Leu"	41..43
FT	unsure		
FT	/tag=	f	
FT	/note=	"encodes Leu"	47..49
FT	unsure		
FT	/tag=	g	
FT	/note=	"encodes Leu"	50..52
FT	unsure		
FT	/tag=	h	
FT	/note=	"encodes Leu"	56..58
FT	unsure		
FT	/tag=	i	
FT	/note=	"encodes Leu"	98..100
FT	unsure		
FT	/tag=	j	
FT	/note=	"encodes Leu"	107..109
FT	unsure		
FT	/tag=	k	
FT	/note=	"encodes Ser"	113..115
FT	unsure		
FT	/tag=	l	
FT	/note=	"encodes Leu"	125..127
FT	unsure		
FT	/tag=	m	
FT	/note=	"encodes Leu"	131..133
FT	unsure		
FT	/tag=	n	
FT	/note=	"encodes Leu"	146..148
FT	unsure		
FT	/tag=	o	
FT	/note=	"encodes Arg"	173..175
FT	unsure		
FT	/tag=	p	
FT	/note=	"encodes Arg"	176..178
FT	unsure		
FT	/tag=	q	
FT	/note=	"encodes Arg"	191..193
FT	unsure		
FT	/tag=	r	
FT	/note=	"encodes Leu"	212..214
FT	unsure		
FT	/tag=	s	
FT	/note=	"encodes Leu"	236..238
FT	unsure		
FT	/tag=	t	
FT	/note=	"encodes Ser"	239..241
FT	unsure		
FT	/tag=	u	
FT	/note=	"encodes Leu"	248..250
FT	unsure		
FT	/tag=	v	
FT	/note=	"encodes Leu"	254..256
FT	unsure		
FT	/tag=	w	
FT	/note=	"encodes Leu"	263..265
FT	unsure		
FT	/tag=	x	
FT	/note=	"encodes Ser"	266..268
FT	unsure		
FT	/tag=	y	
FT	/note=	"encodes Leu"	275..277
FT	unsure		
FT	/tag=	z	
FT	/note=	"encodes Arg"	

FT	unsure		302..304	
FT	/**tag= aa			
FT	/note= "Encodes Ser"			
FT	unsure		311..313	
FT	/**tag= ab			
FT	/note= "Encodes Ser"			
FT	unsure		314..316	
FT	/**tag= ac			
FT	/note= "Encodes Leu"			
FT	unsure		323..325	
FT	/**tag= ad			
FT	/note= "encodes Leu"			
CDS			11..343	
FT	/**tag= ae			
FT	/product= human_preinsulin			
PX	BE-885196-A.			
PD	31-DEC-1980.			
PF	11-SEP-1980; 885196.			
PR	12-SEP-1979; US-075192.			
PA	(REGC) UNIV OF CALIFORNIA.			
DZ	WPI; 81-05762D/05.			
DR	p-psDB; P10053.			
PT	DNA transfer vectors contg. codes for human insulin precursors -			
PT	used to transform microorganisms for insulin producn.			
PS	Claim 4; Page 33; 50pp; French.			
CC	This coding sequence is a generic coding sequence which covers DNA			
CC	sequences able to code for the human preproinsulin amino acid			
CC	sequence. The features table indicates where Leu, Ser or Arg codons			
CC	are intended. These 3 amino acids cannot readily be specified by a			
CC	single generic codon, hence the need for clarification. For each of			
CC	Leu, Arg and Ser the following pairs of codons are what was			
CC	intended, respectively: YTR and CTY (= Leu), MGR and CGY (= Arg)			
CC	and TGN and AGY (= Ser). Transfer vectors containing DNA whose			
CC	sequence corresponds to this generic sequence are claimed.			
CC	See also N10052.			
SQ	Sequence 416 BP;	59 A;	72 C;	85 G; 59 T;
Query Match 11.2%; Score 159; DB 4; Length 416;				
Best Local Similarity 7.6%; Pred. No. 9,55e-02;				
Matches 13; Conservative 56; Mismatches 102; Indels 0; Gaps 0;				
Db	26 mgnytnctcnytntgcncytntngcnynttggcgccncaayccnccngcngcntty	85	: : :	:
	: : :	:	:	:
Cp	277 ARTTYTTCYTGNARNWSNACYTTNWSTGNCRTGTGCRAANCTCYTGNGRGANC	218	:	:
Db	86 gtaaacarcayntgygnwscnayyngdtngargcnyttaytngntngycngar	145	:	:
	:	:	:	:
Cp	217 KNCKNGTNARNGRCAYCGTGNCTCNACRGTGNARNWSYYTTRTCCATNRMGCNKRTANC	158	:	:
Db	146 mrgcgttyttytayaccnaaaracnmgmngmgargcagayyntcar	196	:	:
	:	:	:	:
Cp	157 CRTTKCKKCKRKCAARTACCARCACCDATTARNARNRNACNCCNARDATA	107	:	:
RESULT	6			
ID	N60354 standard; DNA;	231 BP.		
AC	A60354;			
DT	20-JUN-1991 (first entry)			
DE	Desulphatohirudin (I).			
KM	Desulphatohirudin; antibodies; thrombin; ss.			
OS	Synthetic.			
PN	EP-168342-A.			
PD	15-JAN-1986.			
PF	10-JUN-1985; 810268.			

cc times in one h; average body genome; therefore this element provides

1000

cc of the petri

cc of the

PT assays

Cp 353 GGNWSRTANGGNGGNWSYTGTCNGCNWSNARYTTTCTCRANGCNGGNGCRTTN 294

17

Db 114 ngarcarcargngaygaygaraaywsngcngaymgncny 154
| : | : | : : : : | : | : | :
Cp 293 GGNACNAGGTCRCARTTYYTYCTGNARNWSNACYTT 253

Search completed: Tue Aug 27 18:42:40 1996
Job time : 44 secs.

AVAILABILITY

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_tpn n.a. - n.a. Smith-Waterman search, using a protein query
which has been backtranslated into n.a. using IUPAC symbols

Run on: Tue Aug 27 18:43:02 1996; MasPar time 164.99 Seconds
Tabular output not generated.

Title: >US-08-231-565A-2

Description: (1-118) from US08231565A.pep

Perfect Score: 1420

N.A. Sequence: 1 ATGCCNMGNGARGACVGNCA.....SNCCNCCNCTAVNSNCCN 354
Comp: TAGCGNCKNTCTRCNGT.....WNGGNGGNGNATRSWNGN

Scoring table:

Gap 30

Nmatch STD : Dbase 0; Query 0

Searched: 509049 seqs, 177673129 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:STS1 92:STS2
93:STS3 94:STS4 95:STS5 96:STS6 97:STS7
EST-STS-TWO
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100
108:EST101 109:EST110:gnEST2 111:gnEST3 112:gnEST4
113:gnEST5 114:gnEST6 115:gnEST7 116:gnEST8 117:gnEST9
118:gnEST10 119:gnEST11 120:gnEST12 121:gnEST13
122:gnEST14 123:gnEST15 124:gnSTS1 125:gnSTS2 126:enEST1

Database:

EST-STS-TWO
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100
108:EST101 109:gnEST11 110:gnEST2 111:gnEST3 112:gnEST4
113:gnEST5 114:gnEST6 115:gnEST7 116:gnEST8 117:gnEST9
118:gnEST10 119:gnEST11 120:gnEST12 121:gnEST13
122:gnEST14 123:gnEST15 124:gnSTS1 125:gnSTS2 126:enEST1

127:enEST2 128:enEST3 129:enEST4 130:enEST5 131:enEST6
132:enEST7 133:enEST8 134:enEST9 135:enEST10 136:enEST11

Statistics: Mean 61.375; Variance 71.660; scale 0.856

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1392	98.0	593 113	N23770	yx35a12.rl Homo sapie	0.00e+00
2	1392	98.0	593 136	HS770262	yx35a12.rl Homo sapie	0.00e+00
3	1271	89.5	370 117	N35680	yx62h05.rl Homo sapie	0.00e+00
4	1271	89.5	370 135	HS680270	yx62h05.rl Homo sapie	0.00e+00
5	1205	84.9	393 135	HS656271	yx91a09.rl Homo sapie	0.00e+00
6	1205	84.9	393 117	N36656	yx91a09.rl Homo sapie	0.00e+00
7	947	66.7	303 82	T27884	EST19370 Homo sapiens	4.34e-264
8	722	50.8	473 129	HS091255	yx36a12.sl Homo sapie	1.83e-186
9	722	50.8	473 112	N20091	yx36a12.sl Homo sapie	1.83e-186
10	640	45.1	433 135	HS67260	yx91a09.sl Homo sapie	1.17e-158
11	640	45.1	433 113	N24677	yx91a09.sl Homo sapie	1.17e-158
12	540	38.0	396 136	HS796259	yx97f10.sl Homo sapie	2.94e-125
13	540	38.0	396 113	N24796	yx97f10.sl Homo sapie	2.94e-125
14	530	37.3	401 134	HS562267	yx62h05.sl Homo sapie	5.89e-122
15	530	37.3	401 114	N26562	yx62h05.sl Homo sapie	5.89e-122
16	153	10.8	312 129	HS037250	yy15c06.sl Homo sapie	2.17e-09
17	153	10.8	312 109	H94037	16649 Arabidopsis tha	2.17e-09
18	153	10.8	515 72	R90294	16649 Arabidopsis tha	2.17e-09
19	136	9.6	43 117	N36215	yx97f10.rl Homo sapie	1.14e-05
20	136	9.6	43 131	HS215288	yx97f10.rl Homo sapie	1.14e-05
21	132	9.3	411 45	R00378	ye73a04.rl Homo sapie	7.70e-05
22	130	9.2	374 12	H27083	yl16c08.rl Homo sapie	1.97e-04
23	130	9.2	394 12	H26878	yl15e08.rl Homo sapie	1.97e-04
24	130	9.2	479 119	N42998	yy08d01.rl Homo sapie	1.97e-04
25	130	9.2	484 28	H80481	yu75e03.rl Homo sapie	1.97e-04
26	130	9.2	550 12	H27071	yl16a08.rl Homo sapie	1.97e-04
27	128	9.0	572 126	AT35416	18581 Arabidopsis tha	4.97e-04
28	128	9.0	572 117	N37354	18581 Arabidopsis tha	4.97e-04
29	127	8.9	289 33	HSB89F102	H. sapiens partial cd	7.86e-04
30	127	8.9	293 32	HHEA61P	H. sapiens partial cd	7.86e-04
31	126	8.9	338 33	HSBA2A042	H. sapiens partial cd	1.24e-03
32	126	8.9	393 62	RS9572	yg97h10.sl Homo sapie	1.24e-03
33	125	8.8	434 9	H18419	yn49g10.sl Homo sapie	1.95e-03
34	123	8.7	307 42	HUM13601B	Human fetal brain cdn	4.77e-03
35	122	8.6	220 73	R95571	SMNHADA008030SK Schis	7.43e-03
36	122	8.6	274 26	H74486	446 Brassica napus cd	7.43e-03
37	122	8.6	289 38	HSC2PH091	H. sapiens partial cd	7.43e-03
38	122	8.6	300 77	T00011	wEST00732 Caenorhabdi	7.43e-03
39	122	8.6	313 32	HHEF052	H. sapiens partial cd	7.43e-03
40	122	8.6	320 73	R95547	SMNHADA007019T3 Schis	7.43e-03
41	122	8.6	343 22	H62105	yu40f05.rl Homo sapie	7.43e-03
42	122	8.6	396 107	T96010	ye47q01.rl Homo sapie	7.43e-03
43	122	8.6	423 22	H61302	yu39a03.rl Homo sapie	7.43e-03
44	122	8.6	442 107	T95596	ye40b03.sl Homo sapie	7.43e-03
45	122	8.6	528 89	T55327	yb47h10.sl Homo sapie	7.43e-03

ALIGNMENTS

RESULT 1


```

FEATURES
  source
    NCBI gi: 1156822
    Location/Qualifiers
      1..370
        /organism="Homo sapiens"

```

RESULT	4		
ID	HS680270	standard; RNA; EST; 370 BP.	
AC	N35680;		
DT	19-JAN-1996	(Rel. 46, Created)	
DE	19-JAN-1996	(Rel. 46, last updated, Version 1)	
DE	yx62n05.r1	Homo sapiens cDNA clone 266361 5' similar to PIR:A55253	
DE	A55253	melanoma antigen MART-1 - human ;	
KW	EST.		
OS	Homo sapiens	(human)	
OC	Eukaryota;	Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria;	Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.	
NC	[1]		
RP	1-370		
RA	Hillier L., Clark N., Duboue T., Elliston K., Hawkins M.,		
RA	Hollman M., Hultman M., Kucaba T., Le M., Lennon G.,		
RA	Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,		
RA	Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;		
RT	"The WashU-Merck EST Project";		
RL	Unpublished.		
CC	Contact: Wilson RK WashU-Merck EST Project Washington University		
CC	School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,		
CC	MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:		
CC	est@waton.wustl.edu High quality sequence stops: 337 Source: IMAGE		
CC	Consortium, LNL This clone is available royalty-free through LNL		
CC	; contact the IMAGE Consortium (info@image.lnl.gov) for further		
CC	information. NCBI gi: 1156822		
FH	Key	Location/Qualifiers	
FT			
FT	source	1..370	

FT /organism="Homo sapiens"
 FT /clone="266361"
 FT /note="human"
 FT <1..3370
 SQ Sequence 370 BP; 103 A; 87 C; 94 G; 86 T; 0 other;

Query Match 89.5%; Score 1271; DB 135; Length 370;
 Best Local Similarity 58.6%; Pred. No. 0.00e+00;
 Matches 197; Conservative 78; Mismatches 59; Indels 2; Gaps 2;

Db 36 atgccaagaagaatgctcaattcatctatgttaccccaagaagggcgccactct 95
 |||| : ||:|||| ||:||||: ||:|||| ||:|||| ||:|||| ||:||||
 Qy 1 ATGCCWNGARGAYGNCAYTTATHTAYGNTAYCCNARAARGGNCAYGNCAYSN 60
 |||| : ||:|||| ||:||||: ||:|||| ||:|||| ||:|||| ||:||||
 Db 96 tacaccagcgtgaagagcgctggatcgccatcctgacagtgcctggagcttta 155
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Qy 61 TAYACNACNGCARGARGCNGCGNATHGGNATHYTNACNGTNATHYTNNGGNTYN 120
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Db 156 ctgctcatcgctgtgttatgtagaagcaaatggtacagagccttgatgataaa 215
 : || |||| ||:|||| ||:|||| : ||:|||| ||:|||| ||:|||| ||:
 Qy 121 YTNYNATHGCGTGTGCTATGTYGTMGNMGNAAYGNTAYMGNCGNTNATGCAVAAR 180
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Db 216 agttctcatgttgccactcaatgtgc-ttaacaagaagatgcccaagaaggtttgat 274
 : || |||| || ||:|||| ||:|||| : ||:|||| ||:|||| ||:|||| ||:
 Qy 181 WSNYTNCAVYTGNGCNCARTGCGNYTNACNMGNTGYCCNCARGARGNTTYGAY 240
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Db 275 catgggacagcaagtgtctcttcaagagaaaaactgtgaacctgtgggtcccaatgct 334
 |||| |||| : ||:|||| ||:|||| ||:|||| ||:|||| || || ||:||||
 Qy 241 CAYMGNATWSNARGTNWSNTWCARGARAARAAYTGYGRCNGTGTCNCAAYGON 300
 || || || ||:|||| ||:|||| : || ||:||||
 Db 335 ccacctgttatgagaaactctctgcagacagtc 370
 || || || ||:|||| ||:|||| : || ||:||||
 Qy 301 CCNCGCNGTAYG-ARAARYTNWSNGCARGCARMS 335

RESULT 5
 ID HS656271 standard; RNA; EST; 393 BP.
 AC N36656;
 DT 20-JAN-1996 (Rel. 46, Created)
 DT 20-JAN-1996 (Rel. 46, Last updated, Version 1)
 DE yx91a09.r1 Homo sapiens cDNA clone 269080 5' similar to PIR:A55253
 DE A55253 melanoma antigen MART-1 - human i.
 KW EST.
 OS Homo sapiens (human)
 OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
 RN [1]
 RP 1-393
 RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
 RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
 RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
 RA Trevaaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
 RT "The WashU-Merck EST Project";
 RL Unpublished.
 CC Contact: Wilton RK WashU-Merck EST Project Washington University
 CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
 CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
 CC est@wustl.wustl.edu High quality sequence stops: 169 Source: IMAGE
 CC Consortium, LIND This clone is available royalty-free through LIND
 CC ; contact the IMAGE Consortium (info@image.lind.gov) for further
 CC information. NCBI gi: 1157798
 FH Key Location/Qualifiers
 FT source 1..393

FT /organism="Homo sapiens"
 FT /clone="269080"
 FT /note="human"
 FT <1..393
 SQ Sequence 393 BP; 105 A; 94 C; 105 G; 81 T; 8 other;

Query Match 84.9%; Score 1205; DB 135; Length 393;
 Best Local Similarity 56.6%; Pred. No. 0.00e+00;
 Matches 188; Conservative 75; Mismatches 67; Indels 2; Gaps 2;

Db 63 atgccaagaagaatgctcaattcatctatgttncncccaagaagggcgccactct 121
 |||| : ||:|||| ||:||||: ||:|||| ||:|||| ||:|||| ||:||||
 Qy 1 ATGCCWNGARGAYGNCAYTTATHTAYGNTAYCCNARAARGGNCAYGNCAYSN 60
 |||| : ||:|||| ||:||||: ||:|||| ||:|||| ||:|||| ||:||||
 Db 122 tacaccagcgtgaagagcgctggatcgccatcctgacagtgcctggagcttta 181
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Qy 61 TAYACNACNGCARGARGCNGCGNATHGGNATHYTNACNGTNATHYTNNGGNTYN 120
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Db 182 cngctcatcgctgtgttatgtagaagcaaatggtacagagccttgatgataaa 241
 : || |||| ||:|||| ||:|||| : ||:|||| ||:|||| ||:|||| ||:
 Qy 121 YTNYNATHGCGTGTGCTATGTYGTMGNMGNAAYGNTAYMGNCGNTNATGCAVAAR 180
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Db 242 agttctcatgttgccactcaatgtgc-ttaacaagaagatgcccaagaaggtttgat 301
 : || |||| || ||:|||| ||:|||| : ||:|||| ||:|||| ||:|||| ||:
 Qy 181 WSNYTNCAVYTGNGCNCARTGCGNYTNACNMGNTGYCCNCARGARGNTTYGAY 240
 |||| || ||:|||| || || ||:|||| ||:|||| ||:|||| ||:|||| ||
 Db 302 catgggacagcaagtgtctcttcaagagaaaaactgtgnacctgtgggcccacatggg 361
 |||| |||| : ||:|||| ||:|||| ||:|||| ||:|||| || || ||:||||
 Qy 241 CAYMGNATWSNARGTNWSNTWCARGARAARAAYTGYGRCNGTGTCNCAAYGON 300
 || || || ||:|||| ||:|||| : || ||:||||
 Db 362 ccacctgttatgagaaactctctgcagacac 393
 || || || ||:|||| ||:|||| : || ||:||||
 Qy 301 CCNCGCNGTAYGARAARYTNW-SNCGNGARC 331

RESULT 6
 LOCUS N36656 393 bp mRNA EST 16-JAN-1996
 DEFINITION yx91a09.r1 Homo sapiens cDNA clone 269080 5' similar to PIR:A55253
 A55253 melanoma antigen MART-1 - human i.
 ACCESSION N36656
 NID g1157798
 KEYWORDS EST.
 SOURCE human clones-269080 primer=T7 library=Soares melanocyte 2NBHM
 vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
 (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male. 1st strand
 cDNA was primed with a Not I - oligo(dT) primer
 [5'-TCTTACCAATCGACTGGAGCGCGCGACGTTTTTTTTTTTTTTT-3'],
 double-stranded cDNA was size selected, ligated to Eco RI adapters
 (Pharmacia), digested with Not I and cloned into the Not I and Eco
 RI sites of a modified pT73 vector (Pharmacia). Library
 constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
 foreskin melanocytes (fS374) was kindly provided by Dr. Anthony P.
 Albino.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
 Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
 Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 393)
 Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
 Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
 Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
 Trevaaskis E., Waterston R., Williamson A., Wohlmann P. and
 Wilson R.

Db 417 atccatgaagctctgtatccattgctctctacaataccaacagccggtgaagcag 473
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |
Cp 177 RTCCATNAGCNGCKRTANCCRTTNCNCKNGKRCARTACCARCACCCDATNARNAR 121

RESULT 10

ID HS677260 standard; RNA; EST; 433 BP.
AC N24677;
DT 30-DEC-1995 (Rel. 46, Created)
DT 30-DEC-1995 (Rel. 46, Last updated, Version 1)
DE yx91a09.s1 Homo sapiens cDNA clone 269080 3'.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN [1]
RP 1-433
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@watson.wustl.edu High quality sequence stops: 315 Source: IMAGE
CC Consortium, LNL This clone is available royalty-free through LNL
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1138827
FH Key Location/Qualifiers

FT 1..433
FT source /organism="Homo sapiens"
FT /clone="269080"
FT /note="human"
FT mRNA <1..>433
SQ Sequence 433 BP; 119 A; 85 C; 86 G; 142 T; 1 other;

Query Match 45.1%; Score 640; DB 135; Length 433;
Best Local Similarity 56.8%; Pred. No. 1.17e-158;
Matches 108; Conservative 36; Mismatches 45; Indels 1; Gaps 1;

Db 244 ataagtggtgactgtctgcagagagttctcataagcagtgagcattgggaac 303
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |
Cp 348 RTANGNGCNGWSYTGTCNGCNWSNARYTTTCRTANGCNGCNGCRTTNGNAC 289

Db 304 cacaggttcacagttttctctgaagacacacttgcctccgattgatcaaacccct 363
|| ||:||||:||||:||||:|||| |: ||:| | ||:| | ||:| | ||:| | ||:| |
Cp 288 NACNGGYTCRCARTTTTTCYTCGNARNSNACYTTNWSRTCNCKRT-GRTCRAANCCYT 230

Db 364 ctgtgggcacattcttctgaagcacaattgagtgccaacatgaagacttttatcatca 423
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |
Cp 229 CYTNGGRCACNCKNGKNGTARNRCRCAYTGNGTNCNACRTGNARNSYTTTTCATNA 170

Db 424 aggcctgtga 433
: ||:| |
Cp 169 RNCNCCKRTA 160

RESULT 11
LOCUS N24677 433 bp mRNA EST 28-DEC-1995
DEFINITION yx91a09.s1 Homo sapiens cDNA clone 269080 3'.

ACCESSION N24677
NID gi1138827
KEYWORDS EST.
SOURCE human clone-269080 primer=ml3 -40 forward library=Soares melanocyte
2NbmH vector=pf7T3D (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo (dT) primer
[5'-TGTTACCAATCTAGTGGGAGCGCCGACGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pf7T3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 433)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
High quality sequence stops: 315
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1138827
FEATURES Location/Qualifiers
source 1..433
/organism="Homo sapiens"
/clone="269080"
/note="human"

mRNA <1..>433
BASE COUNT 119 a 85 c 86 g 142 t 1 others
ORIGIN

Query Match 45.1%; Score 640; DB 113; Length 433;
Best Local Similarity 56.8%; Pred. No. 1.17e-158;
Matches 108; Conservative 36; Mismatches 45; Indels 1; Gaps 1;

Db 244 ataagtggtgactgtctgcagagagttctcataagcagtgagcattgggaac 303
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |
Cp 348 RTANGNGCNGWSYTGTCNGCNWSNARYTTTCRTANGCNGCNGCRTTNGNAC 289

Db 304 cacaggttcacagttttctctgaagacacacttgcctccgattgatcaaacccct 363
|| ||:||||:||||:||||:|||| |: ||:| | ||:| | ||:| | ||:| | ||:| |
Cp 288 NACNGGYTCRCARTTTTTCYTCGNARNSNACYTTNWSRTCNCKRT-GRTCRAANCCYT 230

Db 364 ctgtgggcacattcttctgaagcacaattgagtgccaacatgaagacttttatcatca 423
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch nn n.a. - n.a. database search, using Smith-Waterman algorithm

```
Run on: Tue Aug 27 08:43:36 1996; MasPar time 1419.91 Seconds
1162.639 Million cell updates/sec
```

Tabular output not generated

Title: >US-08-231-565A-26

Description: (1-2172) from US08231565A.seq

1 GTCAGCGGCATTACCAATC.....AAAAAAAAAAAAAAAAAAAA 2172
 CAGTCGCCGTAATGGTTAG.....TTTTTTTTTTTTTTTTTTTT
 Act Score: 2172
 Sequence:
 Comp:

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0: Query 0

Searched: 270440 seqs. 380027776 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: embl-new3

1:BCT 2:FUN 3:INV1 4:INV2 5:INV3 6:NAM 7:ORG 8:PLN
9:PRI1 10:PRI2 11:PRO 12:ROD 13:SYN 14:UNC 15:VRT
16:VIR1 17:VIR2

Database: genbank92

18:BCRT1 19:BCRT2 20:BCRT3 21:BCRT4 22:BCRT5 23:BCRT6 24:BCRT7
25:BCRT8 26:BCRT9 27:INV1 28:INV2 29:INV3 30:INV4 31:INV5
32:INV6 33:MAM1 34:MAM2 35:PAT1 36:PAT2 37:PAT3 38:PHG
39:PLN1 40:PLN2 41:PLN3 42:PLN4 43:PLN5 44:PLN6 45:PLN7
46:PLN8 47:PRI1 48:PRI2 49:PRI3 50:PRI4 51:PRI5 52:PRI6
53:PRI7 54:PRI8 55:PRI9 56:PRI10 57:PRI11 58:PRI12
59:PRI13 60:ROD1 61:ROD2 62:ROD3 63:ROD4 64:ROD5 65:ROD6
66:ROD7 67:STR 68:SYN 69:UNA 70:VEL1 71:VEL2 72:VEL3
73:VEL4 74:VEL5 75:VEL6 76:VEL7 77:VRT1 78:VRT2 79:VRT3

Database: genbank-new1

80: BCT 81: INV1 82: INV2 83: MAM 84: PHG 85: PLN 86: PRI

87:ROD 88:

Database: u-emb145_9Z

11:11:11

Statistics:
Mean 12.523: Variance 9.406: scale 1.331

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
1	2114	97.3	2130 59	ST3003	gp100-melanocyte line	0.00e+00
2	2094	96.4	2114 55	HMGPM35	Human 95 kD melanocyt	0.00e+00
3	1988	91.5	2026 52	HGU01874	Human me20m mRNA, com	0.00e+00
4	1974	90.9	2131 57	HUMPPEL	Human Pmel 17 mRNA, c	0.00e+00
5	1132	52.1	2817 86	HSPMEL17S2	Human melanocyte-spec	0.00e+00
6	1132	52.1	2817 10	HSPMEL17S2	Human melanocyte-spec	0.00e+00
7	806	37.1	1584 33	BOVRPELA	Bovine retinal pigmen	0.00e+00
8	752	34.6	1881 61	MMU014133	Mus musculus pme117 p	0.00e+00
9	726	33.4	2001 10	HSMELJ03	Human melanocyte prot	0.00e+00
10	726	33.4	2001 86	HSPMELJ03	Human melanocyte prot	0.00e+00
11	245	11.3	1178 10	HSMELJ05	Human melanocyte prot	3.61e-110
12	245	11.3	1178 86	HSPMELJ05	Human melanocyte prot	3.61e-110
13	209	9.6	1559 86	HSPMELJ04	Human melanocyte prot	1.98e-90
14	209	9.6	1559 10	HSMELJ04	Human melanocyte prot	1.98e-90
15	150	6.9	673 86	HSPMELJ02	Human melanocyte prot	1.46e-58
16	150	6.9	673 10	HSMELJ02	Human melanocyte prot	1.46e-58
17	88	4.1	577 10	HSMELJ01	Human melanocyte prot	2.58e-26
18	88	4.1	577 86	HSPMELJ01	Human melanocyte prot	2.58e-26
19	88	4.1	3109 86	HSPMEL17S1	Human melanocyte-spec	2.58e-26
20	88	4.1	3109 10	HSMEL17S1	Human melanocyte-spec	2.58e-26
21	46	2.1	1620 66	RNRNASVNV	R.norvegicus mRNA for	2.21e-06
22	45	2.1	1671 64	RATDBPA	Rat D-binding protein	5.98e-06
23	44	2.0	498 43	S47413	glycine-rich protein	1.60e-05
24	44	2.0	987 40	CLJAC0IV	C.lacryma-jobi mRNA a	1.60e-05
25	44	2.0	1710 39	BNPN12	B.napus (pN12) mRNA f	1.60e-05
26	43	2.0	1934 79	S59519	stannocalcin [Oncoth	4.27e-05
27	43	2.0	2474 42	PEA70HSP	Pisum sativum chlorop	4.27e-05
28	43	2.0	2861 50	HSBINDPR	H.sapiens mRNA for 2	4.27e-05
29	43	2.0	3568 48	HSZ5ABP	H.sapiens mRNA for 2-	4.27e-05
30	42	1.9	517 28	DDE428	Dictyostelium discoid	1.13e-04
31	42	1.9	776 17	CHKTWC	Chicken troponin C (T	1.13e-04
32	42	1.9	736 34	SSCKTITB	Porcine CKII beta mRN	1.13e-04
33	42	1.9	837 59	S78214	APC-tumor suppressor,	1.13e-04
34	42	1.9	1042 59	S61953	c-erbB3-receptor tyro	1.13e-04
35	42	1.9	1305 61	MMMYD118	Mouse mRNA for Myd118	1.13e-04
36	42	1.9	1360 12	MM40808	Mus musculus glutathi	1.13e-04
37	42	1.9	1539 60	CGTUBB3	Cricetulus griseus (c	1.13e-04
38	42	1.9	1714 78	GUSSDBF	C.domesticus mRNA for	1.13e-04
39	42	1.9	1777 30	EMCOLNF13	E.muellieri Lieb. COLN	1.13e-04
40	42	1.9	2062 45	WHTEEFIX	Wheat translation elo	1.13e-04
41	42	1.9	3146 66	RRCPC3145	R.rattus mRNA for pot	1.13e-04
42	42	1.9	3426 81	DRNPAPAT1	D.discoidum mRNA for	1.13e-04
43	42	1.9	3426 5	DRNPAPAT1	D.discoidum mRNA for	1.13e-04
44	42	1.9	3442 12	RNKBSP81P	R.norvegicus mRNA for	1.13e-04
45	42	1.9	3504 29	DNKNR1	Drosophila Kri gene	1.13e-04

ALIGNMENTS

RESULT	1
LOCUS	S73003
DEFINITION	gpl00=melanocyte lineage-specific antigen/Pmel17 homolog [human, mRNA, 2130 nt].
ACCESSION	S73003
NID	g639589
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
PRI	25-JAN-1995

Unclassified.

REFERENCE 1 (bases 1 to 2130)
AUTHORS Adema, G.J., de Boer, A.J., Vogel, A.M., Loonen, W.A. and Figdor, C.G.
TITLE Molecular characterization of the melanocyte lineage-specific antigen gp100
JOURNAL J. Biol. Chem. 269 (31), 20126-20133 (1994)
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gi: 639589] from the original journal article.
COMMENT NCBI gi: 639589
FEATURES
source Location/Qualifiers
1..2130
/organism="Homo sapiens"
/note="human"
22..2007
/note="melanocyte lineage-specific antigen/Pmel17 homolog; melanoma marker protein. Method: conceptual translation supplied by author. NCBI gi: 639590"
/codon_start=1
/product="gp100"
/db_xref="pid:g639590"
/translation="MDLVLRCLLHLAVIGALLAVGATKVPNRQDLGVSRQLRTKAWNRQLYPEWTEAQRDLCHRGQVSLKVSNDPPTLIGANASFTALNFPSCQKVLPGQV
IWNNTILNGSVWGQPVYQETDDACIFPDGGPCPSGWSQKRSFYVWKTWQYQW
QVLGGPVSLSTGTGRAMLTHTEVTVVHRRGSRVYPLAHSSAFTITDQVPSVS
VSQLRALDGNKHFLRNQPLTFALQLHDPSCYLAEADLSYTWDFGDSSTGLISRALV
THTYLEPGVTAQVVAQAIPLTSCGSSPVPTGDGHRPTAEAPNTTAGQVPTVEVG
TTPQAPTAEPSTGTSVPTTEVISTAPVQMPSTAESTGMTPEKVPSEVNGTTLAEM
STPEATMPAEVSIIVLSCTTAQVTTTEWVETTABLELP IPEPEGDASSIMSTESI
TGSIGPLDGTATLRLVKQVPLDCLVLYRGSFSTLDIVQIESAEILQANPSGED
AFELTWSGGPLPKEACMEISSPCQPAQRQLCPVLPSPAQVLVHLKGGSGTYC
LNVSLADTNSLAVYSTOLLMPQOEAGLGVPLVIGILLVMAVLASLIYRRRLMKQD
FSPVQLPHSSHHLRLPFRIFSCPIGENSPFLISGQV"

BASE COUNT 484 a 587 c 575 g 484 t

ORIGIN
Query Match 97.3%; Score 2114; DB 59; Length 2130;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2117; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 11 ggaagacacaatgatcgtgctaaaaagatgcctcttcatttggtcgtgatagggtg 70
|||||
Qy 28 GGAAGACACAATGATCGTGGTCTAAAGATGCGCTTCTTCATTTGGCTGTATAGGTG 87
Db 71 ctttgctggctgggggtacaaaagtacccagaaccaggagactggctggctctcaa 130
|||||
Qy 88 CTTTGCTGGCTGGGGCTACAAAAGTACCAGAAACAGGACTGGCTGGTGTCTCAA 147
Db 131 ggaactcagaacaaagcctggaacaggcagctgtatccagagtggacagaagccaga 190
|||||
Qy 148 GGCACATCAGAACCAAGCCCTGCAAGCAGCAGCTGTATCCAGAGTGGACAGGCCAGA 207
Db 191 gacttgactgctggagaggtggctcaagtgcctcaaggtcagtaatgatggcctacac 250
|||||
Qy 208 GACTTGACTGCTGGAGAGGTGGTCAAGTGTCCCTCAAGGTCAATATATGCGCCCTACAC 267
Db 251 tgattggtgcaaatgcctcctctctattgccttgaacttccctggagcaaaaagttat 310
|||||
Qy 268 TGAATGGTGCAAAATGCCTCCTCTCTATTGCTTGAACCTCCCTGGAAAGCCAAAAGTAT 327
Db 311 tgccagatgggcaggttatctgggtcaacaataccatcatcaatggggccaggtgtggg 370
|||||
Qy 328 TGCCAGATGGCAGGTATCTGGGTCAACAATACCATCATCAATGGGAGCCAGCTGTGG 387
Db 371 gaggagagcagdgatgatccccaggaaactgacgatgcctgcattctccctgatgggtggac 430

Qy 388 GAGGACGACCGATGATACCCAGAGAACTGACGATGCGCTGCATTTCCCTCATGTGTCGAC 447
Db 431 cttgccaatcgtgctctcagaagaagactttgtttatgtctggaagacctggg 490
|||||
Qy 448 CTTGCCATCTGGCTCTTGGTCTCAGAAAGAAAGCTTTTATGCTCGAAGACCTGGG 507
Db 491 gccaatactggcaagtcttagggggccagctgtctgggtcgtgagcattgggacaggcagg 550
|||||
Qy 508 GCAATACTGGCAATTTCTAGGGGGCCCACTGCTGCGCTGACGATTGGGACAGCGCAGG 567
Db 551 caatcgtgggcac 610
|||||
Qy 568 CAATCTGGGCAC 627
Db 611 atgtcctcttgcattccagctcagccttcacattactgacacaggtgcctttctccg 670
|||||
Qy 628 ATGCGCTCTTGTCTATTCGACGCTCAGCCTTACCATTACTGACCAGGTGCTTCTCGG 687
Db 671 tgagcgtgtcccagttggggccttggaaggagggaacacagcacttctgagaatacagc 730
|||||
Qy 688 TGAGGCTGTCCAGTTGGGGCCTTGGATGAGGGGACCAAGCACATTCCTGAGAAATCAGC 747
Db 731 ctctgacetttgccctccagctccatgacccacagtggtctatctggctgaagctgacctct 790
|||||
Qy 748 CTCGTACCTTTGGCTCCAGCTCCATGACCCGAGTGGCTATCTGGCTGAACTGACTCT 807
Db 791 cctacactggacttggagacagtagtggaaacctgatctctcgggacacttggtgtca 850
|||||
Qy 808 CCTACACTGGGACTTTGGACACAGTAGTGGAAACCTGATCTCTCGGCACATTGTGTGCA 867
Db 851 ctcaataactcgtgagcctggccacagctcactgccaggtggtcctgaggtgcacattc 910
|||||
Qy 868 CTCATACTTACCTGGAGCTGGCCAGCTCACTGCCAGGTGCTGCTGCAGGCTGCCATTC 927
Db 911 ctctcacctcgtggtcctcccagttccaggccaccacagatgggcacaggccaactg 970
|||||
Qy 928 CTCCTCACCTCCTGGCTCTCCCGACTTCCAGGCACACAGATGGGCACAGGCCAACTG 987
Db 971 cagaggcccttaacacacacagctggccaaagtgcctactacagaaagtgtgggtactaac 1030
|||||
Qy 988 CAGAGGCCCTTAACACACACAGCTGGCCAAAGTGCCTACTACAGAAAGTTGTGGGTACTACAC 1047
Db 1031 ctggtcaggcgccaaactgcagagacctctggaaccacacatctgtgaggtgccaaacctg 1090
|||||
Qy 1048 CTGGTCAGGGGCCAACTGCAGAGCCCTCTGGAACCAACATCTGTGCAGGTGCCAACCACTG 1107
Db 1091 aagtcataagcaactgcacctgtgcagatgccaactgcagagagcacaggtatgacacctg 1150
|||||
Qy 1108 AAGTCATAAGCACTGCACCTGTGCAGATGCCAACTGCCAGACAGACAGATGACACCTG 1167
Db 1151 agaagtgccagtttcagaggtcatgggtaccacactggcagagatgtcaactccagag 1210
|||||
Qy 1168 AGAAGTGCCAGGTTTCAGAGGTCTAGGGTACCACACTGGCAGAGATGTCAACTCCAGAGG 1227
Db 1211 ctacagatgacacactgcagaggtatccaattgtggtctttctggaaccacagctgac 1270
|||||
Qy 1228 CTACAGGTATGACACCTGCAGAGGTATCAATTGTGTGCTCTTCTGGAACCAACAGCTGCAC 1287
Db 1271 aggtcaacaactacaggtgggtggagaccacagctagagagctaacctatccctgagcctg 1330
|||||
Qy 1288 AGGTAAACAATACAGAGTGGGTGGAACACACAGCTAGAGAGCTACCTATCCCTGACGCTG 1347
Db 1331 aaggtccagatgccaggtcaatcatgtctacggaaagtattacaggttccctggggccccc 1390

QY 1348 AAGGTCAGATGCCAGCTCAATCATGCTACGGAAAGTATTACAGGTTCCTCGGCCCC 1407
|||||
Db 1391 tgcgtgatgtacagccacctaaagctggtagagacaagtccccctggattgtgttc 1450
|||||
QY 1408 TGCTGGATGGTACAGCCACTTAAAGCTGGTGAAGACAGTCCCTCGATGTGTTTC 1467
|||||
Db 1451 tgtatcatatgttctcttcctgcacacctggacatgtccagggtattgaaagtgcg 1510
|||||
QY 1468 TGATCATATGTTCTTTTCGTCACCTGGACATTTGTCAGGATTTGAAAGTGCCG 1527
|||||
Db 1511 agatcctgcaggctgtgccctcggtgaggggatgcatattgagctgactgtgtcctgcc 1570
|||||
QY 1528 AGATCTGCAGGCTGTGCCCTCGGTGAGGGGATGCATTTTGAGCTGACTGTGTCTGCC 1587
|||||
Db 1571 aaggcggtgcacaaagagcctgcattgagatctcatcgcagggtgcccagccctcg 1630
|||||
QY 1588 AAGCGGGCTGCCCAAGCAAGCTGCATGGAGATCTCATCGCCAGGTCGCCAGCCCTG 1647
|||||
Db 1631 cccagcgctgtgccagcctgtgtaccagccagcctgccagctgggtctgcaccaga 1690
|||||
QY 1648 CCCAGCGCTGTGCCAGCTGTGCTACCCAGCCAGCTGCCAGCTGTGTGACCAAG 1707
|||||
Db 1691 tactgaaggtggtcgggacatactgcctcaatgtgtctcgtgctatcaacagcc 1750
|||||
QY 1708 TACTGAAGGTTGGCTCGGGACATATCGCTCAAGTGTCTCTGGTGATACCAAGCC 1767
|||||
Db 1751 tggcagtggtcagcaccagcttcatcagcctggtcgaagagcagccttggcaggttc 1810
|||||
QY 1768 TGGCAGTGTGCAGCACCCAGCTTATCATGCTGCTCAAGAGCAGCGCTTGGCAGGTT 1827
|||||
Db 1811 cgcgtgctggtgggcatcttgcgtgttgatggtgctggtccttgcatctctgatata 1870
|||||
QY 1828 CGCTGATGCTGGGCATCTTGCTGGTTGATGGCTGTGGTCTGTCATCTCTGATATATA 1887
|||||
Db 1871 ggcgcagacttatgaagcaagactctccgtaccagctgtccacatagcagcactc 1930
|||||
QY 1888 GCGCAGACTTATGAAGCAGACTTCTCGTACCCAGCTGCCACATAGCAGAGTCACT 1947
|||||
Db 1931 ggcgtgctacccgcactctgctctgtcccatggtgagaatagcccccctcca 1990
|||||
QY 1948 GGTGCTGTACCCGCACTTCTGCTCTTCTCCATTTGGTGAACAGCCCCCTCTCA 2007
|||||
Db 1991 gtggcgagcaggtctgagctactcatatgatgctgtgatttctcggagtgacagaaa 2050
|||||
QY 2008 GTGGCAGCAGGCTGAGTACTCTCATATGATGCTGTGATTTCTCTGGAGTTGACAGAAA 2067
|||||
Db 2051 cactctattccccagctcttccctggagactactataactgaaataaataactcaga 2110
|||||
QY 2068 CACCTATATTTCCCGCAGCTTCTCCCTGGGAGACTACTATTAACCTGAATAAATACTACA 2127
|||||
Db 2111 gctggaataaaaaaaa 2130
|||||
QY 2128 GCTCAAAAAAATAAAAAA 2147
|||||

RESULT 2
LOCUS HUMPMSS 2114 bp mRNA PRI 15-JUN-1990
DEFINITION Human 95 kD melanocyte-specific secreted glycoprotein mRNA, 3' end.
ACCESSION M32295
NID g183559
KEYWORDS melanocyte-specific secreted glycoprotein.
SOURCE Human melanoma cell line, cDNA to mRNA, clone 8.
ORGANISM Homo sapiens

Eukaryota; Annelida; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Primates; Haplorhini; Catarrhini; Hominiidae.
Vogel, A.
1 (bases 1 to 2114)
Sequence of a melanocyte specific secreted glycoprotein
Unpublished (1990) St. Louis Univ., St. Louis, MO 63104
JOURNAL
COMMENT
Draft entry and computer-readable sequence for [1] kindly submitted
by A.Vogel, 23-FEB-1990.

NCBI gi: 183559
Location/Qualifiers
1..2114
/organism="Homo sapiens"
<1..2028
/note="melanocyte-specific secreted glycoprotein; NCBI
gi: 386754"
/codon_start=1
/db_xref="PID:g386754"
/translation="RGIRKNTMDIVLKRCLLHLAVIGALLAVGATKVRNQDMLGVSR
QLRTKAMNRLYPTENGIRQLQDMRGQVSLKVSNDGPTLIGANAFSIALNFPQSQK
VLPDQVIVWNTIINGSQVWGQPVVPQETDDACIFPDGCPSCWSQKSFVYVW
KTMQYQVQLGCPYSGLSICTGRMLGTHMTVYVYHRRGSRVYLAHSSAFTLD
QYFVSVSQIRALDGNKHFLRNQPLTFALQLHDPGSLAEADLSYTWDFGSSGTL
ISRALVYHTYLEPQVTAQVLIQAAIPLTSCGSSPVGTTDGHRTAEAPNTTAAQV
PTEVVGTPGQAPTAEPSCTTSQVPTTEVISTAPVQMPTEASTMTPEKVPVSEVM
GTTLAEMSTPEATGMTAEVSVIVLSGTAQVTTTETWETARELP IPEPEGRDASS
IMSTETSGSLGLDLDGTALRLVKRQVPLDCVLYRYSFVTLDIVQGESAEILQA
VPSGEGDAFELTVSCGELPKACMEISSPGQPPAORLQVLPSPACOLIVLHOLK
GSGTGYCIANSLADNSLAVSTQLIMPQGEAGLQVPLVIGLILVLAHVILASLIYR
RELKMKQDFVPQLPHSSHLRLPASALLVPLRTAFSSVSRSESHMML"

BASE COUNT 469 a 586 c 575 g 484 t
ORIGIN

Query Match 96.4%; Score 2094; DB 55; Length 2114;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2102; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Db 11 ggaagaacacaaatgagctgggtctaaaaagatgcctctcttcatttggtgctgatatggtg 70
|||||
QY 28 GGAAGAACACAAATGGATCTGGTCTAAAAAGATGCCCTTCTTCATTTGGCTGTGATAGTG 87
|||||
Db 71 ctttgctgctgtgggggtacaaagtaccgaacaccagagactggttgggtgtctcaa 130
|||||
QY 88 CTTTGCTGGCTGTGGGGCTACAAAAGTACCAGAAACCGAGACTGGCTTGGTGTCTCAA 147
|||||
Db 131 ggaactcagaacaaagcctggaacaggcagctgtatccagagtggacagaagcccaga 190
|||||
QY 148 GGAACCTCAGAACAAAGCCCTGGAAACAGGAGCTGTATCCAGAGTGGACAGAGCCACGA 207
|||||
Db 191 gacttgactgctggagaggtggtcaaagtgtccctcaaagtcagtaaatgatgggctaac 250
|||||
QY 208 GACTTGACTGCTGGAGGTGGTCAAGTGTCCCTCAAGGTCAGTAAATGATGATGGGCTACAC 267
|||||
Db 251 tgattggtgcaaatgcctcctctctctattgcttgacttccctggaagccaaaggtat 310
|||||
QY 268 TGATTGGTCAAAATGCCCTCTCTCTATTGGCTTGAATCTCCCTGGAGCCCAAGGTAT 327
|||||
Db 311 tgcagatgggcaggttatctgggtcaacaataaccatcatcaatgggagccaggtggtg 370
|||||
QY 328 TGCAGATGGGCAGGTTATCTGGGTCAACAATACCATCATCAATGGGAGCCAGGCTGGG 387
|||||
Db 371 gagacagcaggtgtatccccaggaactgacgtgcctgcattcttccctgatgggtggac 430
|||||
QY 388 GAGGACACCCAGTGTATCCCCAGAGAACTGACGATGCCCTTCCCTCATGATGGTGGAC 447
|||||

Db 431 cttgccatctggctcttgggtctcagaagaagctttgtttatgtctggaagacctggg 490
|||||
Qy 448 CTTGCCCATCTGGCTCTTGGTCTCAGACAGAGCTTTGTTTATGTCTGGAAGACTGGG 507
|||||
Db 491 gcaataactggcaagtctcagggggcccagtgctcgggctgagcattggagcagcagg 550
|||||
Qy 508 GCCAATACTGGCAATTTCTAGGGGGGCCAGTGTCTGGGCTGAGCATTTGGGACAGGAGG 567
|||||
Db 551 caatctgggcaacacacaccatggaagtgtctacacatcgccggggatccggagct 610
|||||
Qy 568 CAATCTGGGCACACACACCATGGAAGTCACTGTCTACCATCGCGGGGATCCGGAGCT 627
|||||
Db 611 atgtgctcttgcattccagctcagccttcaccattactgaccaggtgctttctccg 670
|||||
Qy 628 ATGTGCTCTTGTCTATTTCAGCTGAGCTTCACCATTTACTGACAGGTGCTTTCTCCG 687
|||||
Db 671 tgagcgtgccagttgcgggaccttgatggaggaacaagcacttccctgagaaatcagc 730
|||||
Qy 688 TGAGGCTGTCCCAGTTGGGGGCTTGGATGAGGGGAACAAGCACTTCCTGAGAAATCAGC 747
|||||
Db 731 ctctgaacttggcctccagctccatgacccccagtggtatctcgtggaactgacctct 790
|||||
Qy 748 CTCTGACCTTTGGCTCCAGCTCCATGACCCAGTGGCTATCTGGCTGAAGCTCACTCT 807
|||||
Db 791 cctacactgggactttgagacagtagtgaacctgatctctcgggcaacttgggtca 850
|||||
Qy 808 CTTACACTGGGACTTTGGAGAGTAGTGGAACTTCTCTGGGCACTTGTGTCA 867
|||||
Db 851 ctcatacttacctgagcctggccagtcactgccaggtggtcctcgaggtgccattc 910
|||||
Qy 868 CTCATACTTACTTGGAGCTTGGGCCAGTCACTGCCAGGTGGTCTGCAAGGTGCCATTC 927
|||||
Db 911 ctctcaactcctgtggtctctcccaagttccaggcaacacagatgggcaaggccaaactg 970
|||||
Qy 928 CTCTCACTCTCTGGCTCTCTCCAGTTCAGGCAACCAAGATGGGCAAGGCAAGCTG 987
|||||
Db 971 cagaggcccttaacacacacagctggccaaagtgcctactacagaagtgtggtgactacac 1030
|||||
Qy 988 CAGAGGCCCTTAACACACAGCTGGCCAAAGTGCCTACTACAGAAAGTTGTGGGTACTAC 1047
|||||
Db 1031 ctgggtcaggcgccaaactcagagacctctggaaccacatctgtgcaggtgccaacactg 1090
|||||
Qy 1048 CTGGTCAAGGCCCACTGCAAGCCCTCTGGAAACCAATCTGTGCAAGGTGGCAACACTG 1107
|||||
Db 1091 aagtcaataagcactgcacctgtgcagatgccaaactcagagagcagaggtatgacacctg 1150
|||||
Qy 1108 AAGTCATAAGCACTGCACCTGTGCAGATGCCAACTGCAGAGAGCAGGATGACACCTG 1167
|||||
Db 1151 agaagtgccagttcagaggtcatgggtaccacactggcagagatgtcaactccagagg 1210
|||||
Qy 1168 AGAAGTGCCAGTTTCAGAGGTCACTGGGTACCACACTGGCAGAGATGTCAACTCCAGAGG 1227
|||||
Db 1211 ctacaggtatgacacctgcagaggtatcaatttggtgctttcttggaaacacagctgcac 1270
|||||
Qy 1228 CTACAGGTATGACACCTGCAGAGGTATCAATTTGTGTGCTTTCTGGAGACACAGCTGCAC 1287
|||||
Db 1271 aggttaacactacagagtggtggagaccacacagctagagagctaccttccctgagcctg 1330
|||||
Qy 1288 AGGTAACTACTACAGTGGGTGGAGCAACAGCTAGAGACTACTATCTCCTGAGGCTG 1347
|||||
Db 1331 aaggtccagatgccagctcaatcatgtctacggaaagtattacaggttccctgggcccc 1390
|||||
Qy 1348 AAGGTCAGATGCCAGCTCAATCATGTCTACGGAAAGTATTACAGGTTCCCTGGGGCCCC 1407
|||||

Db 1391 tctggatggtacagccaccttaagctggtgaagagacagctccccctgattgtgttc 1450
|||||
Qy 1408 TGTGTGATGTGTACAGCCACCTTAAGGCTGTGGAAGACAGACTCCCCCTGGATTTGTTC 1467
|||||
Db 1451 tgtatcagatatggttctctttccgtcacctggacattgtccagaggtattgaaagtgcgc 1510
|||||
Qy 1468 TGTATCGATATGTTCTCTTTTCCGTCACCTTGGACATTTGTCCAGGGGTATTGAAAGTGGCG 1527
|||||
Db 1511 agatcctgcaggtgtgccgtccggtgaggggagtcatttggagctgactgtgctctgcc 1570
|||||
Qy 1528 AGATCTCTCAGAGGCTGTGCCGTCCGGTGAAGGGGATGCATTTTGAGCTGACTGTCTCTCGCC 1587
|||||
Db 1571 aagcgggctgcccaaggaagcctgcattggagatctcatogccaggtgccagccccctg 1630
|||||
Qy 1588 AAGCGGGCTGCCCAAGAGCCCTGCATGGAGATCTCATGCCAGGGTGCCAGCCCTCTG 1647
|||||
Db 1631 ccagcggctgtgccagcctgtgtaccacagccagcctgccagctggttctgcaccaga 1690
|||||
Qy 1648 CCCAGCGGCTGTGCCAGCCTGTCTTACCCAGCCAGCTGCCAGCTGGTTCTGCACCA 1707
|||||
Db 1691 tactgaaaggtggctcggggagacactgctcctcaatgtgtctctggctgataccacagcc 1750
|||||
Qy 1708 TACTCAGAGGCTGGCTCGGGACATATGCTCTCAATGTCTCTCTGGCTGATACCAAGAGCC 1767
|||||
Db 1751 tggcaggtgcagcaccagcttcatgcctggtcaagaagcagccttggcgaggttc 1810
|||||
Qy 1768 TGGCAGTGTGCACACCCAGCTTATCATGCTGTGCTCAGAGCAGCGCCTTGGGCAAGTTC 1827
|||||
Db 1811 cgtcgtgctggcgcacatcctgctggtgtgtgctggtgcttgcattctctgatatata 1870
|||||
Qy 1828 CGTGTATCGTGGGCATCTGTGTGTGTGATGGCTGTGGTCTTGCATCTCTGATATATA 1887
|||||
Db 1871 ggcgcagactatgaagcaagcacttctcgtaccacagttgccacatagcagcagtcact 1930
|||||
Qy 1888 GGGCGACACTTATGAAGCAAGACTTCTCCGTACCCAGTTCGCCACATAGCAGCAGTCACT 1947
|||||
Db 1931 ggcgtgctctaccc-gcatctctgctctgtcccactgttgagagatagccccctctca 1989
|||||
Qy 1948 GGTGCGCTTACCCCGCATCTTCTGCTCTTGTCCCATTTGGTGAGAACAGCCCCCTCTCA 2007
|||||
Db 1990 gtgggcagcaggtgtgagtactctcatgatgtgtgatttctctggagttgacagaaa 2049
|||||
Qy 2008 GTGGGACAGAGTGTGAGTACTCTCATATGATGCTGCTGATTTTCTGGAGTTGACAGAA 2067
|||||
Db 2050 cacctattttccccagctctccctggagactactattaactgaaataaataactcaga 2109
|||||
Qy 2068 CACCTATTATTTCCCCCAGCTCTTCCCTGGGAGACTACTATTAACTGAAATAAATACTCAGA 2127
|||||
Db 2110 gctcg 2114
|||||
Qy 2128 GCCTG 2132
|||||

RESULT 3

LOCUS	HSU01874	2026 bp	mRNA	PRI	27-MAY-1994
DEFINITION	Human me20m mRNA, complete cds.				
ACCESSION	U01874				
NID	q494939				
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;				
	Eutheria; Primates; Catarrhini; Hominiidae; Homo.				

REFERENCE 1 (bases 1 to 2023)
AUTHORS Maresch, G.A., Marken, J.S., Neubauer, M., Aruffo, A., Hellstrom, I.,
Hellstrom, K. and Marquardt, H.
TITLE Cloning and expression of the gene for the Melanoma-Associated ME20
Antigen
JOURNAL DNA Cell Biol. 13, 87-95 (1994)
MEDLINE 94235165
REFERENCE 2 (bases 1 to 2026)
AUTHORS Neubauer, M.G.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-1993) Michael G. Neubauer, Bristol-Myers Squibb
Pharmaceutical Research Institute, 3005 1st Ave, Seattle, WA 98121,
USA

COMMENT NCBI gi: 494939

FEATURES Location/Qualifiers

source

1..2026
/clone="hf12-2"
/clone_lib="cdm8"
/organism="Homo sapiens"
/cell_line="melanoma H3606"
7..78
7..1995
/standard_names="melanoma-associated ME20 antigen"
/note="NCBI gi: 494940"
/codon_start=1
/product="me20m"
/db_xref="PID:g494940"

sig_peptide

CDS

/translation="MDLVLRCLLHVLGALLAVGATKVPKQDLWGLVSRQLRTKAW
NHQLYPDMTEAQRDLDCRGQVSLKVSNDGPTLIIGANASFIALNPFGSKQVLPDQV
IWNVNTIINGSOVWGPVYPQETDDACIFPDGCPGSGWSOKRSFYVWKTWQYW
QVLDGPVSGLSIGTCRAMLGTHMEVTVYHRRGSRSPYLAHSSNAFTIDQVPSVS
VSQLRDLGKNKHFIRNQLITFALQIHDLP SGYLAEDLSYTWDFGDSSGTLISRALV
THYILEPQPVTAQVVLQAAIP LITSCGSPVPGTGDHRPTAEAPNTAGQVPTTVVG
TTPQAPTAEPSCGTSVQVETEVISTAPVQMTAEASTGMTPEKVPVSEVWGTTLAEM
STPEATCMPEVSVIVLSGTAAQVTTTETVETARELP IPEPEGPASSIMSTESI
TGSICPLLDGATLRLVLRQVPLDCLVLYRGSFVTLIDVQIESAEILQAVPSGEGD
AFELTVSCGGLPKPEAKMEIISPQCPAPQALCPVLPSPACOLVHOLIKGSGTTC
LNVSLADNLSLAVSTOLIMPQOEAGGLGQVPLIVGILLVIMAVLIVSLIYRRRLMKQ
DFSPVQLPISHSSHLRLRPLFCSP IGENSEPLISQQV"

mat_peptide

/product="me20m"
BASE COUNT 437 a 564 c 564 g 461 t

ORIGIN

Query Match 91.5%; Score 1988; DB 52; Length 2026;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 2008; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Db 7 atggatcgtgggtctaaagatgcctcttcatttgctgtgataggtgtcttgctggct 66
|||||
Qy 39 ATGGATCTGGTGTGCTTAAAAGATGCCTCTTCATTTGGCTGTGATAGTCTTCTGGCT 98
|||||
Db 67 gtggggctacaaaagtaccagaaaccaggactggcttggtgtctcaaggcaactcaga 126
|||||
Qy 99 GTGGGGCTACAAAAGTACCAGAAACCAGGACTTGGTGTCTCAAGGCCAATCAGA 158
|||||
Db 127 accaaagcttgaaacaggcagctgtatccagagtggacagaagcccgagagacttgactgc 186
|||||
Qy 159 ACCAAAGCTGGAACAGGCAGCTGTATCCAGGTGGACAGAAAGCCCGACAGACTTGCTGC 218
|||||
Db 187 tggagaggttggtcaagtgctcccaagtcagtaagtatgggctacactgattgtgca 246
|||||
Qy 219 TGGAGAGGTGTCAGTGTCCCTCAAGTCACTAAATGATGGGCTACACTGATTGTTGCA 278
|||||

Db 247 aatgcctcctctctatgcttgaacttccctggaagccaaaggtattgccagatggg 306
|||||
Qy 279 AATGCCCTCTCTTATTGCCCTTCAACTTCCTCGAAGCCAAAAGGTATTGCCAGATGGG 338
|||||
Db 307 caggttatctgggtcaacaataaccatcatcaatggagccagggtgtgggaggaagca 366
|||||
Qy 339 CAGGTTATCTGGGTCAACATACCATCATCATGAGGAGCCAGGTGTGGGAGGACGCCA 398
|||||
Db 367 gtatccccaggaaaactgacgatgcctgcatcttccctgatggtgacattgccatct 426
|||||
Qy 399 GTGATATCCCGAGGAAACTGACGATGCGCATCTCCCTGATGTGGACCTTGGCCATCT 458
|||||
Db 427 ggctcttggtctcagaagaagcctttgtttatgtctggaagacctggggccaatactgg 486
|||||
Qy 459 GGCTCTTGCTCTCAGAAGAGAAAGCTTTGTTTATCTCTGGAAGACCTGGGGCCAATCTGG 518
|||||
Db 487 caagttctagggggccagtgcttggtgtgagcattgggacaggcaggggcaatgctgggc 546
|||||
Qy 519 CAATTTCTAGGGGGCCCACTGCTCTGGGCTGAGCATTTGGGACAGCGAGGCCAATGCTGGGC 578
|||||
Db 547 acacacacatggaagtgaactgtctaccatcgccggggtatcccgaggtatgtgctcttt 606
|||||
Qy 579 ACACACACATGGAAGTGACTGTCTACCATCGCGGGGATCCGGAGCTATGTGCTCTT 638
|||||
Db 607 gctcattccagctcagccttcaccattactgaccaggtgctcttctccgtgagcgtgcc 666
|||||
Qy 639 GCTCATTCACGCTCAGCCTTCAACATTAATGACCCAGGTGCTTCTTCCTCGTAGCGGTGCC 698
|||||
Db 667 cagttcggggcttgatggagggaagaacacttctgagaatacagcctctgacettt 726
|||||
Qy 699 CAGTTCGGGGCTTGGATGGAGGGAACAGCACTTCTCTGAGAAATCAGGCTCTGACCTTT 758
|||||
Db 727 qcctccagctccatgacccagtggtctatctggctgaagctgacctctcctacacctgg 786
|||||
Qy 759 GCCCTCCAGCTTCATACCCAGTGGCTATCTGGCTGAAGCTTCTCTCAACCTGG 818
|||||
Db 787 gacttgagacagtagtggaaacctgatctctcggaacctgtggtcactcactaac 846
|||||
Qy 819 GACTTTGGACAGCAGTAGTGAACCCCTGATCTCTCGGGCAGCTTGTGCTCACTCATCTTAC 878
|||||
Db 847 ctggagcctggggcagctcaactgcccaggtggctgagcgtgacctctcctcactccc 906
|||||
Qy 879 CTGGAGCCTGGCCAGCTCACTGCCAGGTGGTCTCTGAGGCTGCCATTCTCTCACTCC 938
|||||
Db 907 tgtggctctcccaagttccaggccacacagatgggacaggcccaactgcagagggccct 966
|||||
Qy 939 TGTGGCTCTCTCCCGACTTCAGGCAACACATGAGGCGACAGGCCCACTGCAGAGGCCCT 998
|||||
Db 967 aacacacacagctggccaaagtgcctactacagaagttgtgggtactacacctggtcagggc 1026
|||||
Qy 999 AACACACAGCTGGCCAAAGTGCCTACTACAGAGTTGTGGGTACTACCTGGTCAAGGG 1058
|||||
Db 1027 ccaactcgagagccctctggaaccacatctgtgaggtggcaaccactgaagtcataagc 1086
|||||
Qy 1059 CCAACTGACAGAGCCCTCTGGAACACATCTCTGACGTTGCCAACCACTGAAGTCTAAGC 1118
|||||
Db 1087 actgcacctgtgagatgccaaactgcagagagcacaggtatgacacctgagaaggtgcc 1146
|||||
Qy 1119 ACTGCACCTGTGCAGATGCCAACTGCAGAGGACAGAGTATGACACCTCAGAGGTTGCCA 1178
|||||
Db 1147 gtttcagaggtcatgggtaccacactggcagagatgtcaactccagaggtacaggtatg 1206
|||||
Qy 1179 GTTTCAGAGGTCAATGGGTACCACTTGGCCAGAGATCTCACTCCAGAGGCTACAGGTATG 1238
|||||

Db 1207 acaccgtcagaggtatacaattggtgtcttcttggaaccacagctgcacaggtacaact 1266
|||||
Qy 1239 ACACCTGCAGAGGTATCAATTGTGTGCTTTCGGAACACAGCTGCACAGTAACAAC 1298
|||||
Db 1267 acagagtggttgagaccacagctagagctacctatccctgagcctgaaggtccagat 1326
|||||
Qy 1299 ACAGAGTGGGTGGAGCACACAGCTAGACGCTACCTATCCCTGAGCCTGAAGTCCAGAT 1358
|||||
Db 1327 gccagctcaatcatctacggaagaattacaggttccctgggcccctctggtggt 1386
|||||
Qy 1359 GCCAGCTCAATCATGTCTACGGAAGATTATACAGGTTCCCTGGGCCCCCTCTCGATGCT 1418
|||||
Db 1387 acagccaccttaagctggtgaagagacaagctccccctcgattgtttctgtatgatgat 1446
|||||
Qy 1419 ACAGCCACCTTAAGCTGCTGAAGACACAACTCCCTCGGATGTTCTCTATCGCATAT 1478
|||||
Db 1447 ggtcccttttcogtccacctggacattgccaaggtattgaagtgccgagatcctgcag 1506
|||||
Qy 1479 GGTTCCTTTTCGTCGACCTGGACATGCTCGAGGTTATGAAGTCCGAGATCTGCAG 1538
|||||
Db 1507 gctgtgccgtccggtgaggggatgattgagctgactgtcctgccaaagcggtgtg 1566
|||||
Qy 1539 GCTGTGCCGTCCGGGTGAGGGGATGCAATTTGAGCTGACTGTGTCTGCCAAGCGGGCTG 1598
|||||
Db 1567 ccaaggaagctcatgagatctcatgcagaggtccagcccctgccagcggtgtg 1626
|||||
Qy 1599 CCCAAGGACGCTCATGTCAGATCTCATGCCAGGTCGCCAGCCCTGCCACGGGCTG 1658
|||||
Db 1627 tccagctgtgctacccagccagcctgccagctggtctgacacagatcgaaggt 1686
|||||
Qy 1659 TGCCAGCCTGTCTACCCAGCCAGCTGCCAGCTGTTCTGCACACAGATCTCAAGGT 1718
|||||
Db 1687 ggctcggggacatactgctcaatgtctctgtggtgataccacagcctggcagtggtc 1746
|||||
Qy 1719 GCCTGGGGACATACGCTCAATGTCTCTGGCTGATACCAACAGCCTGGCAGTGGTC 1778
|||||
Db 1747 agcaccagcttatcatgctggtcaagaagcaggggccttgggcaggttccgctgac 1806
|||||
Qy 1779 AGACCCAGCTTATCATGCTGTGTCOAAGACAGG---CCTTGGGCGAGTTCGCTGATC 1835
|||||
Db 1807 gtgggcattctggtgttgatggtctggtccttgcatctctgatatagggccaga 1866
|||||
Qy 1836 GTGGGCATCTTCTGCTTGTGCTGTGCTTGTGCTTGTGCTATATATAGGCGCAG 1895
|||||
Db 1867 ctatgaagcaagcttctcgtacccagttgccacatagcagtcactgggtgcgt 1926
|||||
Qy 1896 CTTATGAGCAAGACTTCTCCGTACCCAGTTGCCACATAGCAGCAGTCACTGGCTGGCT 1955
|||||
Db 1927 ctaccocgcatctctgctcttgccttggagaaatagccccctcctcagtgggcag 1986
|||||
Qy 1956 CTACCCGCGCATCTTCTGCTTCTGCTTGGTGGTGAGACAGCCCCCTCTCAGTGGGCGAG 2015
|||||
Db 1987 caggtctgagtactctcatatgatgtgatt 2019
|||||
Qy 2016 CAGGTCGAGTACTCTCATATGATGCTGTGATT 2048
|||||

RESULT 4
LOCUS HUMPEL 2131 bp mRNA PRI 08-JAN-1995
DEFINITION Human Pmel 17 mRNA, complete cds.
ACCESSION M77348
NID q190105
KEYWORDS Pmel 17 protein; melanocyte.
SOURCE Homo sapiens skin cDNA to mRNA.

ORGANISM Homo sapiens
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
REFERENCE 1 (bases 1 to 2131)
AUTHORS Kwon,B.S., Chintamaneni,C., Kozak,C.A., Copeland,N.G.,
Gilbert,D.J., Jenkins,N., Barton,D., Francke,U., Kobayashi,Y. and
Kim,K.K.
TITLE A melanocyte-specific gene, Pmel 17, maps near the silver coat
color locus on mouse chromosome 10 and is in a syntenic region on
human chromosome 12
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (20), 9228-9232 (1991)
MEDLINE 92021023
COMMENT NCBI gi: 190105
FEATURES
source
location/Qualifiers
1..2131
/organism="Homo sapiens"
/cell_line="primary culture"
/cell_type="melanocyte"
/sequenced_mol="cDNA to mRNA"
/tissue_type="skin"
1..2121
/gene="Pmel 17"
12..2018
/gene="Pmel 17"
/note="NCBI gi: 190106"
/codon_start=1
/product="Pmel 17 protein"
/db_xref="PID:g190106"
/translation="MDILVKRCLLHLAVIGALLAVGATKVPKQDMLGVSRLRTKAM
INQLPTEWAQRDLDRGQVQVSLKVNDSGPTLIGANASFSIALNFPQSKVLPDQGV
INWNTITINGSQWQGPVPTQDADACIFPDGCFPSGWSQKRSFYVMKWTQGYW
QVLGSPVSGLSICTGRAMLGTHTMEVTVYHRRGSRVYPLAHSSAFTITDQVPFSVS
VSQRLADGKNKHLRNQPLTEALQLHDPFSGYLAERADLSYTWDFGDSSTGLISRAPVV
THYLEPQVTAQVVLQAAIPLTSCSSPVPTGTDHRTPAEPNTTAQGVPTTEVG
TTPQAPTAERSCTTSVOVPTTEVISTAPVOMPTAESTGMTPEKVPVSEVMGTIAEM
STPEATGTPAEVSVIVLGGTTAAQVTTETWETARELP IPEPEGDASSINSTESI
TGSGLPDDGTATLRVKKRQVLDVLYRGSFVTLDIVQGESAEILQAVPSEGD
AFELTVSCGGPKPEACMEISSPQPPAQRQVLPSPAQVLVHLQILKGSCTYC
LNVSADTNSLAVVSTQLIMPVPGILLTQGEAGLQVRLIVGLILVMAVVLASLIYR
RLMLKQDFSVLPQLPFSSSHWLRLPRIFCSPIGENSPLISGQQV"

BASE COUNT 475 a 588 c 578 g 490 t
ORIGIN

Query Match 90.9%; Score 1974; DB 57; Length 2131;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 2105; Conservative 0; Mismatches 5; Indels 21; Gaps 1;
Db 1 ggaagacacacatgatctggtgtctaaaaagatgccttcttcttcttctgctgtagagtg 60
|||||
Qy 28 GGAAGACACACATGCACTCTGCTGTCTAAAAAGATGCTCTTTCATTGGCTGTGATAGTG 87
|||||
Db 61 cttgtgctgctgtgggggtacaaagtaccacgaacacagcagctggttgggtgtctcaa 120
|||||
Qy 88 CTTTCTGCTGTGGGGCTTACAAAAGTACCAGAAACACGACTGGCTGGTGTCTCAA 147
|||||
Db 121 ggcacactcagaacacaaagcctggaacagcagctgtatccagagtggaacagaagccaga 180
|||||
Qy 148 GGCAACTCAGAACCAAGCCTGGAAACAGGAGCTGTATTCACAGAGTGGACAGAGCCACGA 207
|||||
Db 181 gacttgactgctgagaggtggtcaagtgctccctcaagtcagtaaatgagggcctaac 240
|||||
Qy 208 GACTTCACTGCTGAGAGTGTCAAGTGTCCCTCAAGGTCAGTAATGATGAGGCTACAC 267
|||||
Db 241 tgattggtgcaaatgcctcctctctctattgcttgaaactcctctggaagcacaaggtat 300
|||||

QY 268 TGATTGGTCAAAATGCGCTCTCTCTATTGGCTTGAACTTCCCTGGAAGCAAAAAGGTAT 327
DB 301 tqccagatgggcaggttatctgggtcaacaataaccatcatcaatggagccaggtgtggg 360
QY 328 TCCAGATGGGCAGGTTATCTGGGTCAACAAATACCATCATCAATGGGAGCCAGGTGGG 387
DB 361 gaggacagccagtgatcccccaggaaactgaagatgacctctctccatgagtggtggac 420
QY 388 GAGGACAGCCAGCTGTATCCCCAGGAATACTCAGCATGCCCTGCATCTCTCCCTGATGGTGGAC 447
DB 421 cttgcccatctggctcttggctcagagaagaagctttgttatgtctcgtgaagacctggg 480
QY 448 CTTGCCATCTGGCTCTTGGCTCTCAGAGAGAAAGCTTTGTTATGCTCGAAGACCTGGG 507
DB 481 gccaatactggcaagttctagggggcccagtgctggggtgagcattgggagcaggcagg 540
QY 508 GCCAATACTGGCAATTTCTAGGGGGCCCACTGCTGGGCTGAGCATTTGGGACAGGACGG 567
DB 541 caatgctgggcacacacacacatggaagtgactgtctaccatgccggggggtatcccagct 600
QY 568 CAATGCTGGGCACACACACATGGAAGTGACTGTCTACCATGCCGGGGATCCCGAGCT 627
DB 601 atgtgctcttgcctcattccagctcagccttcaccattactgaccagtgctcttctccg 660
QY 628 ATGTGCTCTTCTGCTATTCAGCTCAGCCTTACCATTTACTGACCAGGTGCTTCTCCG 687
DB 661 tqagcgtgccaggttgcgggcttgatggagggaacaagaacctctcagaaatcagc 720
QY 688 TGAAGCTGTCCAGTTGCGGGCTTGGATGGAGGGAACAAAGCACTTCTCGAANAATCAG 747
DB 721 ctctgaccttggccctccagctccatgacctagtggctatctggctgaagctgacctct 780
QY 748 CTCGTGACCTTTGGCTTCAGCTCCATGACCCAGTGGCTATCTGGCTGAAGCTGACCTCT 807
DB 781 cctaaccttgggaacttggagacagtagtggaaacctgatctctcgggcaacctgtgtca 840
QY 808 CCTACACTGGGACTTTGGAGACACTAGTGGAAACCTGATCTCTGGGCACCTTGTGGTCA 867
DB 841 ctcatcttacctgagcctggccagtcactgcccaggtggtcctcgagctgcccattc 900
QY 868 CTCATACTTACCTGGAGCCTGGCCCACTGCTGCCAGGTGGCTTGCAGGCTGCCATT 927
DB 901 ctctcaactctgtgctcctccccagttccaggccaccacagatgggcaagcccaactg 960
QY 928 CTCTCACTCTGTGGCTCTCCCACTTCCAGGCAACAGATGGGACAGGCAAGCTG 987
DB 961 cagaggccctaaacacacagctggccaagtgcctactacagaagtgtgggtactaac 1020
QY 988 CAGAGGCCCCCTAACACACACAGCTGGCCAACTGCTACTACAGAAGTTGTGGGTACTAC 1047
DB 1021 ctgggtcagggcccaactgcagagccctctggaaccacatctgtgcagtgccaacctg 1080
QY 1048 CTGGTCAAGGCGCAACTGCAGAGCCCTCTGGAAACCATCTGTGCAGGTGCCAACCACTG 1107
DB 1081 aagtcataagcactgacctgtgcagatgccaaactgcagagacaggtatgaaacctg 1140
QY 1108 AAGTCAATAGCACTGCACCTGTGCAGATGCGCAACTGCAGAGACAGGATATGACACCTG 1167
DB 1141 aqaaggtgccagttccagaggtcatgggtaccacactggcagagatgtcaactccagagg 1200
QY 1168 AGAAGTCCCACTTTTCAGAGGCTCATGGCTACCACACTGGCAGAGATGTCAACTCCAGAG 1227
DB 1201 ctacaaggtatgacacctgcagaggtatcaattgtgggtcttcttggaaaccacagctgca 1260

QY 1228 CTACAGGTATGACACTGCAGAGGTATCAATTGTGTGCTTTCTGGAACACAGCTGCAC 1287
DB 1261 aggttaacaactacagagtgggtggagaccacagctagagagtaacctatccctggagcctg 1320
QY 1288 AGGTAACTACAGACTGGGTGGAGACACAGCTAGAGAGCTTACCTATCCCTGAGCCTG 1347
DB 1321 aaggtccagatgccagctcaatctctacggaagtattacaggttccctggggccccc 1380
QY 1348 AGGTCCAGATGCCAGCTCAATCTGTACGGAAGTATTACAGGTTCCCTGGGCCCCC 1407
DB 1381 tgcctggatggtacagccacacttaagctggtgaagacacaagtccccctgattgtgttc 1440
QY 1408 TGCTGATGTGTACAGCCACTTAAGGCTGGTGAAGACAAGTCCCTCGATTGTGTTC 1467
DB 1441 tctatcatatgggttcttccgtcacccctggacattgtccaagggatatgaaagtgcg 1500
QY 1468 TGTATCGATATGTTCCCTTTTCCGTCACTCCCTGCACATTTGTCCAGGCTATTGAAAGTGCG 1527
DB 1501 agatcctgcaggtgtgcogtccggtgagggggatgcatttgagctgactgtctcctgcc 1560
QY 1528 AGATCTGCAGGCTGTGCCCTGCCGTGAGGGGGATGCATTTGAGCTGACTGTCTCTGCC 1587
DB 1561 aaggggggtgcccagaagcctgcattgagatctcatcgccaggtgcccagcccccctg 1620
QY 1588 AGGGGGGGCTGCCCAAGAGAGCTGCATGGAGATCTCATGCCAGGGTGCACGCCCTCG 1647
DB 1621 cccagcggctgtgccagcctgtctaccagcccagctgccagctggtctgcaccaga 1680
QY 1648 CCCAGCGGCTGTGCCAGCCTGTGCTTACCAGCCAGCTGCCAGCTGGTTCTGCCACAGA 1707
DB 1681 tactgaaggggtgggtcggggacatactgctcaatgtgctctggctgataccaacagcc 1740
QY 1708 TACTCAAGGTTGGCTCGGGGACATACTGCCCTCAATGTCTCTGTGCTGATACCAACAGCC 1767
DB 1741 tggcaggtggcagcaccagcttatcatgctgtcctgggattcttctcacaggtcaag 1800
QY 1768 TGGCAGTGGTGCAGACCCAGCTTATCATGCTG-----GTCAAG 1806
DB 1801 aagcagccttgggcaggttcgggtgatcgtgggcatcttgggtgtgatgggtgtgg 1860
QY 1807 AAGCAGGCTTGGGCAGGTTCCGCTGATTCGTGGGCATCTTGTGTGTGATGGCTGTGG 1866
DB 1861 tccctgcatctctgatataataggcgagacttatgaagcaagacttctccgtaccaccagt 1920
QY 1867 TCCTTGGATCTCTGATATATAGGCGCAGACTTATGAGCAAGACTTCTCCGTACCCCACT 1926
DB 1921 tqccaataagcagcagtcactgggtgogtctaccocgactcttctgtcttgccttcattg 1980
QY 1927 TCCCACTATGACAGCTCAGCTGGCTGGCTGTACCCCGCATCTTCTGCTCTTGTGCCATTG 1986
DB 1981 gtgaagatagccccctcctcagtgggcagcaggtctgagtaactctcatatgatgtgtga 2040
QY 1987 GTGACAGACAGCCCCCTCTCTCAGTGGGAGCAGGCTGTGAGTACTCTCATATGATGCTGTGA 2046
DB 2041 ttttctcggagttgacagaacacatatttccccagctcttccccggagagactactat 2100
QY 2047 TTTTCTCGGAGTTGACAGAAACACCTTATTTTCCCCCAGCTTTTCCCTGGGAGACTACTAT 2106
DB 2101 taactgaataaataactcagagcctgaaaaa 2131
QY 2107 TAACCTGAATAAATACTCAGAGCCTGAAGAAA 2137

Aug 27 08:56

US-08-231-565A-26.rgs

15

RESULT 5
LOCUS HSPMEL1752 2817 bp DNA PRI 01-JAN-1996
DEFINITION Human melanocyte-specific (pmel 17) gene, exons 2-5, and complete cds.
ACCESSION U20093
NID gl142634
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; eukaryote crown group; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonotata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2817)
AUTHORS Kwon,B.S., Kim,K., Heng,H.H., Shi,X.M., Tsui,L., Lee,Z.H., Yoon,B. and Pickard,R.T.
TITLE Genomic organization and FISH mapping of human pmel 17, the putative silver locus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2817)
AUTHORS Kwon,B.S.
TITLE Direct Submission
JOURNAL Submitted (05-JAN-1995) Byoung S. Kwon, Indiana University School of Medicine, Microbiology and Immunology, 635 Barnhill Drive, Indianapolis, IN 46202, USA
COMMENT NCBI gi: 1142634
FEATURES
source
location/Qualifiers
1..2817
/clone lib="lambda FIX II from Stratagene"
/organism="Homo sapiens"
/map="12pter-q21"
/chromosome="12"
join (U19491:1111..1186,67..459,853..1014,1266..2396,2499..2722)
CDS
/gene="pmel 17"
/note="NCBI gi: 1142636"
/codon start=1
/product="Pmel 17"
/db_xref="PDB:1l42636"
/translation="NDLVKRCCLHLAVIGALLAVGATKVPKNDQWLVGSRLRTKAW
NRQLYPTEAQRIDCWGQVSLKVSNDGPTLIGANASFSIALNFPQSKVLPDQV
IWNNTIINGSQVWGCQVPYQETDDACIFPDGCPSCGSQKRSFVYVWKTWQYV
VQLGDPVSLGSGTGRAMLGHTEVTVTHRRGSRVPLAHSSAFITDQVPSVS
VSQLALDGNKHFILRNQPLTFALQHPDPSGLAEADLSYTWDFGDSGLISRAPV
THYLEPQVTAQVVLQAAIPLTSCSSPVPCTTDGHRTEAEPNTAQVPTTEVG
TPGQAPTEPSTGVSQVPTTEVISTAPVQMPTEKVPVSEVMGTTIADM
STPEATGMFAEYSIVLSGTTAAQVTTTEWVTTARELPPEPEGPDASSIMSTESI
TCSLPLDGTATILWRQVPDQVLYRGSFVTLDIVOGIESAEILQAVPSGCD
AFELTVSGGQKPEACMEISSPGCPQAPRLQVPLPSPAQLVHLQILGSGGTYTC
LNVSLADNSIAVAVSTQLIMPQGEAGLQVPLIVGLILMAVVLASLIYRRLLKQD
FSPVPLPHSSSHWLRLPRIFGSCP IGENSPLLSGQV"
67..459
/gene="pmel 17"
/number=2
853..1014
/gene="pmel 17"
/number=3
1266..2396
/gene="pmel 17"
/number=4
2499..>2722
/gene="pmel 17"

Aug 27 08:56

US-08-231-565A-26.rgs

16

BASE COUNT 627 a 772 c 739 g 679 t
ORIGIN
/number=5
Query Match 52.1%; Score 1132; DB 86; Length 2817;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1265 gaccaggtgcttttccctgagcgtgtcccagttgcggccttggatggagggaacaag 1324
Qy 669 GACCAAGTGCCTTTCTCCGTGAGCGTGTCCAGTTGCGGCGCTTGATGAGGGAACAAG 728
Db 1325 caattctgagaatacagcctctgacctttgcccctccagctccatgacccagtggtat 1384
Qy 729 CACTTCTCAGAAATCAGGCTGTGACCTTTGCCCTCCAGCTCCATGACCCAGTGGCTAT 788
Db 1385 ctggtggaagctgacctctctacacacctgggactttggagacagtagtgaaccctgac 1444
Qy 789 CTGGCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACAGTAGTGAACCCCTGATC 848
Db 1445 tctgggcaacctgtgtgctcaactcatacttaactggagcctggcccagtcactgcccagtg 1504
Qy 849 TCTGGGCACTTGTGTCTACTCATCTTACTTGAGGCTGGCCAGTCACTGCCAGGTG 908
Db 1505 gtctcgagctgacctctctcacctcctgtggtcctcccagttccaggcaccaca 1564
Qy 909 GTCTCGAGGCTGCCATTCTCTCACCTCTGTGGCTCTCCGCCAGTTCCAGGCAACACA 968
Db 1565 gatgggcaagcccaactgcagagggccctaacaccacagctggccaagtgcctactaca 1624
Qy 969 GATGGGCACAGCCCACTGCAGAGGCCCTTAACACCACTGCGCAAGTGCCCTACTACA 1028
Db 1625 gaagttgtgggtactacacctgtcaggcgccaaactgcagagccctctggaaccacatct 1684
Qy 1029 GAAGTTGTGGGTACTACCTGCTGTCAGGCCCACTGCAGAGCCCTCTGGNACCACATCT 1088
Db 1685 gtcgaggtgccaaacctgaagtcataagcactgcacctgtgcagatgccaaactgcagag 1744
Qy 1089 GTGCAGGTGCCAACCACTGAAGTCATAAGCACTGCACCTGTGCAGATGCCCACTGCAG 1148
Db 1745 aqcacaggtatgacacctgagaagtgccagtttcagaggtcatgggtaccacactggca 1804
Qy 1149 ACCACAGGTATGACACCTGAGAAGTGCCAGTTTCAGAGGTATGGGTACCACTGGCA 1208
Db 1805 gagatgtcaactccagaggctacaggtatgacacctgcagagaggtatcaattgtggtctt 1864
Qy 1209 GAGATGTCAACTCCAGAGGCTACAGGTATGACACCTGCAGAGGTATCAATTGTGGTCTT 1268
Db 1865 tctggaaccacagctgcacaggtacaactacagagtggtgtggagaccacagctagagag 1924
Qy 1269 TCTGGAACCAACAGTGCACAGGTAACTACAGAGTGGGTGGAGACCACTAGCTAGAG 1328
Db 1925 ctacstatacctgagcctgaaggtccagatgccagctcaatcatgtctacggaaagtatt 1984
Qy 1329 CTACCTATCCCTTGAGGCTGAGAGTCCAGATGCCAGCTCAATCATCTACGGAAGTATT 1388
Db 1985 acaggttccctgggcccctgctgagatggtacagccaccttaagctggtgaagagaca 2044
Qy 1389 ACAGGTTCCCTGGGCCCTGCTGGATGGTATGAGCCACCTTAAGGCTGGTGAAGACAA 1448
Db 2045 gtccccctggattgtctctgctatgctatggttcttcttcctgcacacctggacattgctc 2104
Qy 1449 GTCCCCCTGGATTGTCTCTATCGATATGCTTCTTTCCTTTTCCGTCACCCCTGGACATTGTC 1508

Db 2105 cagggtattgaaagtcgccagatcctgcaggctgtgccgtccggtgaggggatcattt 2164
|||||
Qy 1509 CAGGGATTGAAGTCGCCAGATCCTCGAGGCTGTGCGCTCGCGTGAGGGGATGCAATT 1568
|||||
Db 2165 gacgtgaetgtgtcctgccaaaggcggtgccaaaggaagcctgcattgagatctcatcg 2224
|||||
Qy 1569 CAGCTGACTGTGTCTGTCGCCAAGCGGGCTGCCAAGAGAGCCTGCTGAGATCTCATCG 1628
|||||
Db 2225 ccagggtgccagccctcccaagcgctgtccagcctgtgctaccagccagcctgc 2284
|||||
Qy 1629 CCAGGGTGCACGCCCTCCACAGCGCTGTCCAGCCTGTCTACCCAGCCAGCCTGC 1688
|||||
Db 2285 cagctggttcgaccagatcactgaagtggtcggtcggggaactactcctcaatgtgtct 2344
|||||
Qy 1689 CAGCTGCTTTCACACAGATACTGAAGGCTGCTCGGGGACATACGCTCAATGTCT 1748
|||||
Db 2345 ctgctgataccacagcctggcagtggtcagcaccagccttatcattgcctggt 2398
|||||
Qy 1749 CTGCTGATACACAGCCTGGCAGTGTTCAGCACCAGCCTTATCATGCTGCT 1802
|||||

RESULT. 6

ID HSMELI752 standard; DNA; PRI; 2817 BP.
AC U20093;
DT 01-JAN-1996 (Rel. 46, Created)
DT 01-JAN-1996 (Rel. 46, Last updated, Version 1)
DE Human melanocyte-specific (pmel 17) gene, exons 2-5, and complete cds.

OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.

RN 1-2817

RA Kwon B.S., Kim K., Heng H.H., Shi X.M., Teui L., Lee Z.H.,

RA Youn B., Pickard R.T.;

RT "Genomic organization and FISH mapping of human pmel 17, the

RT putative silver locus";

RL Unpublished.

RN 1-2817

RP 1-2817

RA Kwon B.S.;

RT ;

RL Submitted (05-JAN-1995) to the EMBL/GenBank/DBJ databases.

RL Byoung S. Kwon, Indiana University School of Medicine, Microbiology

RL and Immunology, 635 Barnhill Drive, Indianapolis, IN 46202, USA

CC NCBI gi: 1142634

FB Key

Location/Qualifiers

FT source

FT 1..2817

FT /clone_lib="Lambda FIX II from Stratagene"

FT /organism="Homo sapiens"

FT /chromosome="12"

FT /map="12pter-q21"

FT join(019491:1111..1186,67..459,853..1014,1266..2396,2499..

FT 2722)

FT /gene="pmel 17"

FT /codon_start=1

FT /product="Pmel 17"

FT /db_xref="PDB:gi1142636"

FT /translation="MDLVKRLHLAVIGALLAVAGTKVPRNODMLGVSRQLRTKAW

FT N

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT VNNIINGSQVWGQVVPQETDDACIFPDGCPSCSHSQKRSFYVYWKTHGQVQV
L
FT GGPVSGLSITGTRAMLGTHMEVTVYHRRGSRSVPLAHSSSAFTITDQVPFVSVSQ
L
FT RALDGNKHFLRNQPLITFALQLHDPSCYLAEDALSYTMDFGDSSGTLISRAPVWTHY
L
FT EPGVTAQVVLQAAIPLISCGSPVPCTTDGHRPTAEAPNTTAGOVPTTEVVGTPCQ
A
FT PTAEPSCGTTSVQVPTTEVISTAPVQMPATAESTGMPTEKVPVSEVMGTTLAEMSTPEAT
G
FT MTPAEVSIWLSGTTAAQVTTTEWETTRARELP IPEPEGPDASSIMSTESITGSLGPL
L
FT DGTATLRUVKQVPLDQVLYRYGFSFVTLIDIVQGISAEILQAVPSGEGDAFELLVSC
Q
FT GGLPKACMETISSPGQPAPQRLQCPVLPSPAQLVHLQILKGGSGTYCIANSLADTN
S
FT LAVVSTQLIMPQGEAGLQVPLIVGILVLMVAVLASLIYRRRLMKQDFVSPQLPHSS
S
FT HMLRPRIFCSCPIGENSPILLSQQV*
FT 67..459
FT /gene="pmel 17"
FT /number=2
FT 853..1014
FT /gene="pmel 17"
FT /number=3
FT 1266..2396
FT /gene="pmel 17"
FT /number=4
FT 2499..>2722
FT /gene="pmel 17"
FT /number=5
FT
FT
SQ Sequence 2817 BP; 627 A; 772 C; 739 G; 679 T; 0 other;

Query Match 52.1%; Score 1132; DB 10; Length 2817;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1265 gaccaggtgccttctcgtgagcgtgtccagttgcggccttgatgagggagacaag 1324
|||||
Qy 669 GACCAGTGCCTTTCTCGTGAGGCTGCCAGTTGGGGCTTGGATGAGGGAACAAG 728
|||||
Db 1325 cacttctgagaaatcagcctctgaccttgccctccagctccatgacccagtgctat 1384
|||||
Qy 729 CACTTCTCAGAAATCAGCTCTGACCTTGGCCCTCCAGCTCCATGACCCAGTGCTAT 788
|||||
Db 1385 ctgctgaagctgacctctctacacctggactttggagacagttagtgaacctgac 1444
|||||
Qy 789 CTGGCTGAAGCTGACCTCTCTCACA CCGTGGACTTTGGAGACAGTAGTGAACCTGATC 848
|||||
Db 1445 tctcgggcactgtggtcactcatactactcagcctggagcctggccagtcactgccaggtg 1504
|||||
Qy 849 TCTCGGCACTTGTGTGCTACTCATATTTACTCTGGAGCTGGCCAGTCACTGCCAGGTG 908
|||||
Db 1505 gtctcgaggtgacctctcactcctgtggtctctccagttccagggcaccaca 1564
|||||
Qy 909 GTCTCGAGGCTGCCATCTCTCTCA CCGTCTGGCTCTCTCCAGTTCCAGGCAACCACA 968
|||||
Db 1565 gatggcacaggcccaactgcagagggccctaaacacacagctggccaagtgcctactaca 1624
|||||
Qy 969 GATGGGCAGAGCCCACTGCAGAGCCCTTACACACACAGCTGGCCAGCTTACTTACA 1028
|||||
Db 1625 gaagtgtgggtactacacctggtcaggcgcccaactgcagagccctctggaaccacatct 1684
|||||
Qy 1029 GAAGTTGTGGGTACTACACTGGTGCAGGCGCAACTGCAGACCCCTCTGGAACACATCT 1088
|||||

Db 1685 gtccaggtgccaccactgaatgataaagcactgcacactgtgcagatgccaactgcagag 1744
|||||
Qy 1089 GTCCAGGTGCCACCACCTCAAGTCATTAAGCACTGCACCTGTCCAGATGCCAACCTGCAGAG 1148
Db 1745 agcacaggtatgacacctgaagaagtgccaggttcagaggtcatggtgataccacactggca 1804
|||||
Qy 1149 AGCACAGGTATGACACCTGAGAGGTGCCAGTTTCAGAGGTCTATGGTATACCACTGGCA 1208
Db 1805 gagatgtcaactccagaggtcacaggtatgaacactgcagaggtatcaattggtgctt 1864
|||||
Qy 1209 GAGATGTCAACTCCAGAGGTACAGGTATGACACCTGCAGAGGTATCAATTGTGTGCTT 1268
Db 1865 tctggaaaccacagctgcacaggtacaactacaggtgggtggagaccacagctagagag 1924
|||||
Qy 1269 TC7GGAAACCACAGCTGCACAGGTAACTACAGAGTGGGTGGACACCACTAGAGAG 1328
Db 1925 ctacctatccctgagcctgaaggtccagatgccaggtcaatcatgtctacggaaagtatt 1984
|||||
Qy 1329 CTACCTATCCCTGAGCCTGAAGGTCCAGATGCCAGCTCAATCATGTCTACGGAAAGTATT 1388
Db 1985 acaggttccctgggcccctgctggatggtacagcaccctaaagctggtggaagacaa 2044
|||||
Qy 1389 ACAGGTTCCCTGGGCCCTGCTGGATGGGTACAGCCACCTTAAAGCTGCTGAAGACAA 1448
Db 2045 gtccccctggattgtgtctgtatcagataggttcccttccgtccaccctggacattgtc 2104
|||||
Qy 1449 GTCCCCCTGGATTGTCTGTATGATATGTTCTTTTCGCTCACCCCTGGACATTGTC 1508
Db 2105 cagggtattgaaagtcagagactcctgcaggtgtgcgtccggtgaggggagtcattt 2164
|||||
Qy 1509 CAGGTTATTGAAGTGCACAGATCCTGCGAGGTGTGCCGTCCGCTGAGGGGATCGATTT 1568
Db 2165 gactgactgtctctgccaaagcggtgccaaaggaagcctgcattgagatctcatcg 2224
|||||
Qy 1569 GAGCTGACTGTCTCTGCCAAGGGGGGCTGCCAAGGAAGCCTGCATGAGATCTCATCG 1628
Db 2225 ccaggggtgcaagccccctgccagcggtgtgccagcctgtgctaccagccagcctgc 2284
|||||
Qy 1629 CCAGGGTGCACAGCCCTGCCACGGCTGTGCCAGCCTGTGCCAGCTGTGTACCCAGCCAGCCTGC 1688
Db 2285 cagctggttgcaccagataactgaagggtggtcggggacataactgcctcaatgtct 2344
|||||
Qy 1689 CAGCTGTTCTGCACAGATACTGAAGGTGGCTCGGGGACATACTGCCTCAATGTGTCT 1748
Db 2345 ctggctgataccacagctggcagtggtcagcaccagccttatcagcctggt 2398
|||||
Qy 1749 CTGGCTGATACCAACAGCCTGGCAGTGGTGCAGACCCAGCCTTATCATGCTGCTGTT 1802
|||||

RESULT 7 BOVBRPE1A 1584 bp mRNA MAM 08-MAR-1993

LOCUS Bovine retinal pigment (RPE1) mRNA, 3' end.

DEFINITION M81193

ACCESSION gl62770

NID retinal pigment.

KEYWORDS Bos taurus (Library: Unizap XR) calf retinal pigment epithelium

SOURCE cDNA to mRNA.

ORGANISM Bos taurus

Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;

Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae;

Bovini.

REFERENCE 1 (bases 1 to 1584)

AUTHORS Kim, R. Y. and Wistow, G. J.

TITLE The cDNA RPE1 and monoclonal antibody HMB-50 define gene products preferentially expressed in retinal pigment epithelium

JOURNAL Curr. Eye Res. 55, 657-662 (1992)

COMMENT NCBI gi: 162770

FEATURES Location/Qualifiers

source 1..1584

/organism="Bos taurus"

/dev_stage="calf"

/sequenced_mol="cDNA to mRNA"

/tissue_type="retinal pigment epithelium"

/tissue_lib="Unizap XR"

1..1477

/partial

/gene="RPE1"

/note="ORF; NCBI gi: 162771"

/codon_start=2

/product="retinal pigment"

/db_xref="PID:gl62771"

/translations="OYQVVLGGPVSGLSIGTDKAMLGTYNMEVTYHRRGSGSYVPLA
HSSAFTTDDQVFFSVSVSQALDGRNKRFLKQPLTFALQLHDPSCYLADLSYT
WDFGDTCTLISALVTHTYLESQVTAQVVIQAAIPLTSGSPVPCTDRHVTTA
EAPGTTAGQVPTTEVMGTTTGPVTAEPGTTTCWVPTTDDVGTTPQVAVTSKVLST
PVEMPTAKATGRTPVEVSTTESGTTVQTTPPELVETAGVSTPEPAGSNTSSFHPT
ETAGLSPLPDDTATLVLEKRAPLDICVLVYGSFSLDIDVSIETAEIQAQVSSE
GDAFELTVSCGGGLPKEACMDISSPCOLPAORLCQVPPSPACQLAHLVHLKGGST
YCLNLSLADANSIAMVSTQLVMPGQENGRLQAPLFGVILLVLTALLASLIYRRILMK
QSEVPLPQLPHGRTQWLRLPWFVRSCTPESKPLLSQQV"

BASE COUNT 340 a 473 c 418 g 353 t

ORIGIN

Query Match 37.1%; Score 806; DB 33; Length 1584;

Best Local Similarity 82.7%; Pred. No. 0.00e+00;

Matches 1348; Conservative 0; Mismatches 230; Indels 52; Gaps 21;

Db 1 ccaataactggcaagtctctggggggccagtgctcgactgaagcattcggaagcagcaagc 60

Qy 509 CCAATACTGGCAATTTCTAGGGGGCCAGTGTCTGGCTGAGCATTTGGACAGGAGGC 568

Db 61 aatgctgggcaacataaacatggaagtgaagtgactgtctaccaccggggggtccacagacta 120

Qy 569 AATGCTGGGCACACACACCATCGAAATGACTGTCTACCATCGCGGGGATCCCGAGGCTA 628

Db 121 tgtgccccctcgctcaactccagttcagccttcaccattactgaccagtgccctctctgt 180

Qy 629 TGTGCTCTTCTGCTATTCAGCTCAGCCTTACCATTTACTGACAGGCTGCCTTTCCTGCT 688

Db 181 gagt-gt-gt-cagctgcagctgcagcgttgatggaaggaacaagcctctctgagaaagcagcc 240

Qy 689 GAGGCTGCCAGTTGCGGGGCTTGAGTAGAGGGAACAAGACACTTCTTGAGAAATCAGCC 748

Db 241 tctgaccttgccctccagctccatgatccagtggtctatttgctgggggtgaccttc 300

Qy 749 TCTGACCTTTGCTCCCTCCAGCTCCATGACCCCACTGGCTATCTGGCTGAGCTGACCTCTC 808

Db 301 ctacacctgggaatttggtgacagtacagggaccctgatctctcgtgggcaactcaagctcac 360

Qy 809 CTACACCTGGGACTTTGGAGACAGTACTGGAACCTGATCTCTCGGGCACCTTGTGCTCAC 868

Db 361 tcacacttacctagagtctggcccagtcactgcaggtgcaggtggtgctgcaggtgaccttc 420

Qy 869 TCATACTTACCTGGAGCCTGGGCCCACTGCTGCTGCCAGCTGGCTGCGAGGCTGCCATTC 928

Db 421 tctcacctcctgtggtcctctccaggtccaggtccagagcactacagatgagcagctgcacactgc 480

|||||

Qy	1889	GGCAGCTTATGAGCAGACT-----TCTCCGTACCCCACTTGCACATAGCAGCAG	1942
Db	1396	ccagtggctaagctctgccctgggtgtctccgctcttgcgccattggtgagagcaaacccct	1455
Qy	1943	TCACATGGCTGGCTACCCCGCATCTTCTGCTCTTGTCCCATTTGGTGAAGAACGCCCTC	2002
Db	1456	cctcagtgagacacaggtctgaagtcgtcttatgtgaagtcatgatttaccag-gtgac	1514
Qy	2003	CTTCAGTGGGCAGCAGGCTCAGTACTCTCATATGATGCTGTGATTTCTCGAGTTGAC	2122
Db	1515	agcaaggcctgtctttctgtgtcttccttcagagactaccattgcctgaataaaagac	1574
Qy	2063	AGAAACACTATATTTCCCCAGCTCTCCCTGGGAGACTACTATTAACTGAATAAATAC	2122
Db	1575	tcagaacttg 1584	
Qy	2123	TCAGAGCCTG 2132	
RESULT	8		
LOCUS	MMU14133	1881 bp	mRNA
DEFINITION	Mus musculus pmel17 protein mRNA, complete cds.	ROD	07-JUL-1995
ACCESSION	U14133		
NTID	9887940		
KEYWORDS	.		
SOURCE	mouse.		
ORGANISM	Mus musculus		
REFERENCE	1 (sites)		
AUTHORS	Kwon, B.S., Halaban, R., Ponnazhagan, S., Kim, K., Chintamaneni, C., Bennett, D. and Pickard, R.T.		
TITLE	Mouse silver mutation is caused by a single base insertion in the putative cytoplasmic domain of Pmel 17		
JOURNAL	Nucleic Acids Res. 23, 154-158 (1995)		
MEDLINE	95175358		
REFERENCE	2 (bases 1 to 1881)		
AUTHORS	Kwon, B.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-AUG-1994) Byoung S. Kwon, Indiana University School of Medicine, Microbiology and Immunology, 635 Barnhill Dr, Indianapolis, IN 46202, USA		
COMMENT	NCBI gi: 887940		
FEATURES	Location/Qualifiers		
source	1..1881		
	/clone="pmel17-10"		
	/strain="c57Bl/6"		
	/organism="Mus musculus"		
	/map="chromosome 10 si locus"		
	/chromosome="10"		
	/sex="female"		
	/cell_line="primary culture"		
	/cell_type="melanocyte"		
	/tissue_type="skin"		
	1..1881		
	/gene="pmel17"		
	/note="NCBI gi: 887941"		
	/codon_start=1		
	/product="pmel17 protein"		
	/db_xref="PID:g887941"		
	/translation="MVGVRRSFLPVLVLSALLAVGAGLEGRNODWLGVPRLVTKTW		
	NRLQYPWFTEVCGSNCRGQVSLRVINDGPTLVGANASFIALHFGCSKVLPGQV		
	IWANNVTINGSQWVGQVPYDPEDDCAVFPDGGCPSPGSRKPPPSFVYVWTKTWKTC		
CDS			

Qy 716 TGGAGGGAACAAGCACTTCTCTGAGAAATCAGCCTCTGACCTTTGCCCTCCAGCTCCATGA 775

FTT
/note="DSEG number: D12S53E"

RESULT	9
ID	standard; DNA; PRI; 2001 BP.
U35ME103	
AC	U31797;
DT	23-DEC-1995 (Rel. 46, Created)
DT	23-DEC-1995 (Rel. 46, Last updated, Version 1)
DE	Human melanocyte protein Pmel 17 gene, exons 4, 5, and 6.
KW	.
OS	Homo sapiens (human)
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
OT	[1]
RP	1-2001
RA	Bailin T., Lee S.T., Spritz R.A.;
RT	"Genomic organization and sequence of D12S53E (Pmel 17), the human
RT	homologue of the mouse silver (sl) locus";
RL	Unpublished.
RL	[2]
RP	1-2001
RA	Bailin T.;
RT	;
RL	Submitted (17-JUL-1995) to the EMBL/GenBank/DBJ databases.
RL	Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of
RL	Wisconsin, Madison, WI 53706, USA
CC	NCBI gi: 1125059
Key	Location/Qualifiers
FT	source
FT	1..2001
FT	/organism="Homo sapiens"
FT	/chromosome="12"
FT	/map="12ql3-q14"
FT	/note="DSEQ number: D12S53E"

FT exon 95..229
/number=4
FT exon 623..781
/number=5
FT exon 1030..1755
/number=6
SQ Sequence 2001 BP; 457 A; 554 C; 492 G; 498 T; 0 other;

Query Match 33.4%; Score 726; DB 10; Length 2001;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1030 gaccaggtgcctttcccgtagagcgtgtccacgttggccttgcctccagctccatgacccagtggtat 1149
|||||
Qy 669 GACCAGGTGCCTTTCCTCGAGCGGTGCCAGTTGGCGGCTTGGATGAGGGAACAAG 728
|||||
Db 1090 cacttctgagaatcagcctctgaccttggcctccagctccatgacccagtggtat 1149
|||||
Qy 729 CACTTCTGAGAAATCAGCCTCTCACCTTGGCTCCAGTCCATGACCCAGTGGCTAT 788
|||||
Db 1150 ctggtggaagctgacctctcctacactgggacttggagacagtgtggaacctgac 1209
|||||
Qy 789 CTGGCTGAAGCTGACCTCTCTCACCTGGGACTTTGGAGACAGTAGTGGAAACCTGATC 848
|||||
Db 1210 tctcgggcacttggteactcatacttacctgagcctggccagtcactgccaggtg 1269
|||||
Qy 849 TCTCGGGCAGCTTGTGGTCTACTACTTACCTTGGAGCGCTGGCCAGTCACTGCCAGGTG 908
|||||
Db 1270 gtctcggagcgtgcattctcctcactcctgtgtgtcctcctcagctggccagtcactgccaggtg 1329
|||||
Qy 909 GTCTCGAGGTGCGCAATTCCTCTCACTCTCTGGCTCTCCCAAGTTCCAGGACACCA 968
|||||
Db 1330 gatggcacagcgaactgacagagcgcctaacacacagctggccagtgctactaca 1389
|||||
Qy 969 GATGGGACAGGCGCAATTCAGAGGCGCCCTAACACACAGCTGGCGAAGTGCCTACTACA 1028
|||||
Db 1390 gaagttgtgggtactacacctggctcaggcgccaactgcagagccctctggaacacatct 1449
|||||
Qy 1029 GAAGTTGTGGGTACTACCTCTGTCAGGCGCCAACTGCAGAGCCCTCTGGAACCACTCT 1088
|||||
Db 1450 gtgcaggtgccaacacactgaagtcataaagcactgcaactgtgcagatgccaactgcagag 1509
|||||
Qy 1089 GTGCAGGTGCCAACCACTGAAGTCATAAGCACTGCACCTGCAGATGCCAACTGCAGAG 1148
|||||
Db 1510 agcacaggtatgacacctgagaaggtgccagtttcagaggtcatgggtaccacactggca 1569
|||||
Qy 1149 AGCACAGGTATGACACTGACGAGGTGCCAGTTTCAGAGGTATGGGTGACCACTGGCA 1208
|||||
Db 1570 gagatgtcaactccagaggtacaggtatgacacctgcagaggtatcaattgtggtcctt 1629
|||||
Qy 1209 GAGATGTCAACTCCAGAGCTACAGGTATGACACCTTCGACAGGTATCAATTGTGTGCTT 1268
|||||
Db 1630 tctggaacacagctgcacaggttaacaactacagagtgggtggagaccacagctagagag 1689
|||||
Qy 1269 TCTGGAACCACTGCACAGGTAAACACTACAGAGTGGGTGGAGACCACTAGAGAG 1328
|||||
Db 1690 ctacctatccttgagcctgaaggtccagatgccaggtccaatcatgtctacggaaagtatt 1749
|||||
Qy 1329 CTACCTATCTCCCTGAAGCTTCAAGTCCAGATGCCAGTCAATCATCTCTCGGAAGTATT 1388
|||||
Db 1750 acaggt 1755
|||||
Qy 1389 ACAGGT 1394
|||||

RESULT 10
LOCUS HSPMEL03 2001 bp DNA PRI 19-DEC-1995
DEFINITION Human melanocyte protein Pmel 17 gene, exons 4, 5, and 6.
ACCESSION U31797
NID g1125059
KEYWORDS .
SEGMENT 3 of 5
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2001)
AUTHORS Bailin, T., Lee, S.-T. and Spritz, R.A.
TITLE Genomic organization and sequence of D12S53E (Pmel 17), the human homologue of the mouse silver (si) locus
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 2001)
AUTHORS Bailin, T.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of Wisconsin, Madison, WI 53706, USA
COMMENT NCBI gi: 1125059
FEATURES
source
1..2001
/organism="Homo sapiens"
/chromosome="12"
/map="12q13-q14"
/note="DSEG number: D12S53E"
95..229
/number=4
623..781
/number=5
1030..1755
/number=6
BASE COUNT 457 a 554 c 492 g 498 t
ORIGIN

Query Match 33.4%; Score 726; DB 86; Length 2001;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1030 gaccaggtgcctttcccgtagagcgtgtccacgttggccttggatggaggaacaag 1089
|||||
Qy 669 GACCAGGTGCCTTTCCTCGAGCGGTGCCAGTTGGCGGCTTGGATGAGGGAACAAG 728
|||||
Db 1090 cacttctgagaatcagcctctgaccttggcctccagctccatgacccagtggtat 1149
|||||
Qy 729 CACTTCTGAGAAATCAGCCTCTGACCTTTGGCTCCAGTCCATGACCCAGTGGCTAT 788
|||||
Db 1150 ctggtggaagctgacctctcctacactgggacttggagacagtgtggaacctgac 1209
|||||
Qy 789 CTGGCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACAGTAGTGGAAACCTGATC 848
|||||
Db 1210 tctcgggcacttggteactcatacttacctgagcctggccagtcactgccaggtg 1269
|||||
Qy 849 TCTCGGGCAGCTTGTGGTCTACTACTTACCTTGGAGCGCTGGCCAGTCACTGCCAGGTG 908
|||||
Db 1270 gtctcggagcgtgcattctcctcactcctgtgtgtcctcctcagctggccagtcactgccaggtg 1329
|||||
Qy 909 GTCTCGAGGTGCGCAATTCCTCTCACTCTCTGGCTCTCCCAAGTTCCAGGACACCA 968
|||||
Db 1330 gatggcacagcgaactgacagagcgcctaacacacagctggccagtgctactaca 1389
|||||

QY	969	GATGGGCACAGCCAACTGACGAGAGGCCCTTACACACAGCTGGCCAGTGCCTACTACA	1028
Db	1390	gaagtttgtggtactacacctggtcaggcgccaactgcagagccctctggaaccacatct	1449
QY	1029	GAAGTTGTGGGTACTACACCTTGCTCAGCGGCCAACTGCAGAGGCCCTCTGGAACACATCT	1088
Db	1450	gtcagtggtgccaaaccaactgaagtcataagcaactgcaactgtgcagatgccaaactgcagag	1509
QY	1089	GTCCAGGTGCCAACCACTGAAGTCAATAGCACTGCACCTGTGCAGATGCCAACTGCAGAG	1148
Db	1510	agcacaggtatgacacctgagaagtgccagdttcagaggtcattcaggtaccacactgcga	1569
QY	1149	AGCACAGTATGACACCTGCAAGAGGTGCCATTTTCAGAGGTCTGGGTACCACACTGGCA	1208
Db	1570	gagatgtcaactccagaggtacagdtatgacacctgcagaggtatcaattgttgtgctt	1629
QY	1209	GAGATGTCMACTCCAGAGGCTACAGGTATGACACCTGCAGAGGTTACAAATTGTGGTGCTT	1268
Db	1630	tctggaaccacagctgcacaggtgaacaaactacagagtgggtggagaccacagctagagag	1689
QY	1269	TCTGGACCAACAGCTGCACAGGTTAACAACTACACAGCTGGGTGGAGACCAACAGCTAGAGAG	1328
Db	1690	ctacctatccctgagcgtgaaggtccagatgccagctcaatctgtctacggaaagtatt	1749
QY	1329	CTACCTATCCCTGAGCCTGAAGTCCGATGCCAGCTCAATCATGTCTACGGAAAGTATT	1388
Db	1750	acaggt	1755
QY	1389	ACAGT	1394

RESULT	11	
ID	HSMEI05	standard; DNA; PRI; 1178 BP.
AC	U31799;	
DT	23-DEC-1995 (Rel. 46, Created)	
DT	23-DEC-1995 (Rel. 46, Last updated, Version 1)	
DE	Human melanocyte protein Pmel 17 gene, exon 11 and partial cds.	
KW	.	
OS	Homo sapiens (human)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.	
RP	{1}	
RP	1-1178	
RA	Bailin T., Lee S.T., Spritz R.A.;	
RT	"Genomic organization and sequence of D1Z53E (Pmel 17), the human	
RT	homologue of the mouse silver (si) locus";	
RL	Unpublished.	
RN	{2}	
RN	1-1178	
RA	Bailin T.;	
RT	;	
RT	Submitted (17-JUL-1995) to the EMBL/GenBank/DBPJ databases.	
RL	Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of	
RL	Wisconsin, Madison, WI 53706, USA	
CC	NCBI gi: 1125061	
EH	Key	Location/Qualifiers
FT	source	1..1178
FT		/organism="Homo sapiens"
FT		/chromosome="12"
FT		/map="12q13-q14"
FT		/note="DSEG number: D1Z53E"
FT	CDS	join(U31808:158..232,U31807:128..241,U31807:400..551)

2

FT :95..229,U31797:623..781,U31797:1030..1755,U31798:382..494
 FT U31798:736..821,U31798:932..1135,U31798:1238..1327,298..43
)
 FT /note="gp100; ME20; NCBI gi: 1125063"
 FT /codon_start=1
 FT /product="melanocyte protein Pmel 17"
 FT /db_xref="PID:gi1125063"
 FT /translation="MDLVKRCILHLAVTIGALLAVGATKVPKQNDMLGVSQRLTKAW
 N RQLYPEWTEAQRLLDCHRGQVSKVSNDCPTLIGANASFIALNFPCKQKVLPGQV
 W VNNTINGSQWGCQVPVPEQTDACIPDGGPCPGSGWSQKRSFYVWKTWGQYHQV
 L GCPVSGLSICTGRAMLGTHTEVTVVYHRGSRSVYPLAHSSSAFTITDQVPFVSYSVQ
 L RALDGGNKHFIRNQPLTFALQLHDPGSGYLAEADLSYTWDFGDSGTLISRALVWTHY
 FT EPGVTAQVVLQAAIPLTSCGSSVPVPTDGHRTAEAPNTTAGQVPTTEVVGTTGQ
 A PTAEPSCTTSHQVPTTEVISTAPVQMPPTAESTGCTKPKVPSEVMGCTTIAEMSTPEAT
 G MTPAEVSIWLSGTAAQVTTTEWVTTARELPIPEPEGPDASSIMSTESITGSLGPL
 L DGTATIRILWKQVPLDCLVLYRGSGFVTLDIQVIGESAEILQAVPSGEGDAFELTVSC
 Q GCLPKENAEISSPGCPQPARLQCPVPSAPCOLVHLQILKGSCTYCLNVSLLADTN
 FT LAVVSTQLIMPQCEAGLGQVPLVIGILLVMAVVASLIYRRLRMKQDFSVPLPHSS
 S HMLRLPRIFCSGPIGNSPLLSGQV"
 SQ Sequence 1178 BP; 298 A; 231 C; 373 G; 276 T; 0 other;

Query Match	11.38;	Score 245;	DB 10;	Length 1178;
Best Local Similarity	99.6%;	Pred. No. 3.61e-110;		
Matches 246;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
Db	295	agggcgcagctatgaagcaagactcttcggtaccgccagttgccacatagcagcagtcac	354	
Qy	1887	AGGCGCAGCTATTGAAGCAAGACTTCGCGTACCAGCTGGCCATAGCAGCAGCTCAC	1946	
Db	355	tggtcgogtctaccccgcatcttctgtcttgccttcccatggtgagatagagccctcttc	414	
Qy	1947	TGCGTGGCTCTACGCGGCATCTTGTCTCTTGTCCATTGGTGAGACAGCGCCCTCCTC	2006	
Db	415	agtgggcagcaggtctgagtactctcatatgatgctgattctctggagttgcagaa	474	
Qy	2007	AGTGGGCAGCAGGCTGAGTACTCTCATATGATGCTGTGATTTCTCTGGAGTTGACAA	2066	
Db	475	acacctatttccccagctctccctgggagactactattaaactgaataaataactcag	534	
Qy	2067	ACACCTATATTCCCCCAGCTTCCTCTGGGAGACTACTATTAACTGAAATAAATACTCAG	2126	
Db	535	agcctga	541	
Qy	2127	AGCCTGA	2133	
RESULT	12			
LOCUS	HSPMEL05	1178 bp	DNA	PRI - 19-DEC-1995
DEFINITION	Human melanocyte protein Pmel 17 gene, exon 11 and partial cds.			
ACCESSION	U31799			
Q1125061				
NID				

Aug 27 08:56

US-08-231-565A-26.rgc

29

KEYWORDS
SEGMENT 5 of 5
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1178)
TITLE Baillin, T., Lee, S.-T. and Spritz, R.A.
JOURNAL Genomic organization and sequence of D12553E (Pmel 17), the human
homologue of the mouse silver (si) locus
AUTHORS Unpublished (1995)
TITLE 2 (bases 1 to 1178)
JOURNAL Baillin, T.
AUTHORS Direct Submission
TITLE Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445
JOURNAL Henry Mall, University of Wisconsin, Madison, WI 53706, USA
COMMENT NCBI gi: 1125061
FEATURES
source
1..1178
/organism="Homo sapiens"
/chromosome="12"
/map="12q13-q14"
/note="DSEG number: D12553E"
Join(U31808:158..232,U31807:128..241,U31807:400..546,
U31797:95..229,U31797:623..781,U31797:1030..1755,
U31798:382..496,U31798:736..821,U31798:932..1135,
U31798:1238..1327,298..432)
/note="gpl00; ME20; NCBI gi: 1125063"
/codon_start=1
/product="melanocyte protein Pmel 17"
/db_xref="PID:g1125063"
/translation="MDLVKRLCLHLAVIGALLAVGATKVPKQMLGVSRQLRTKAW
NRQYPTWTEAQRLDQWRGVSLKNSDPTLIGANNASFSIALNFFGSRQVLPDQV
IWNQNTINGQVWGQVPYPOETDADICFPDGGPCPSGWSOKRSFYVVKWTGQVY
QVLGCPVSGLSIGTCRAMLGHTMEVTVHRRGSRVYPLAHSSAFTIDQVPSVS
VSQRALDGNKHE LRNQPLTFALQDLHAEADLSYTWDFGDSGTLISRALVY
THYLEPGEVTAQVVLQAAIPLTSCGSPVPTDGHRTAEAPNTTAGQVPTTEWG
TTPQAPTAEP SGTSVQVPTTEVI STAPVQMP TAESTCMTEKVPVSEVMTTIAEM
STPEATGMPEAENSIVLSGTTAAQVTTTEWVETARELP IPEEGPDASSIMTESI
TGSIGPLDLGTATLRVLRQVPLDQVLYRGSFVTDIVOGIESAEILQAVPSEGD
AFELTVSCQGLPKEACMEISSPCQPPAQRQCQVLPSPACQLVLIHQILKGGSTYC
LNVSLADTNSLVAVTQLMPEQEGALGGQVPLIVGILLVIMAVLASLIYRRLMKQD
FSVPQLPHSSHMLRPLRFICSCPIGENSPILSGQV"

CDS

BASE COUNT 298 a 231 c 373 g 276 t
ORIGIN
Query Match 11.3%; Score 245; DB 86; Length 1178;
Best Local Similarity 99.6%; Pred. No. 3.61e-110;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 295 aggcagacattatgaagaagacttctccgtaccacagttgccacatagcagcgtcac 354
Qy 1887 AGCGGACAGATTATGAAGAAGACTTCTCCGTACCACGTTGCCACATAGCAGTCAC 1946
Db 355 tggctgcgtaccacagcattcttctgttctgttccattggtgagaatagccccccttc 414
Qy 1947 TGGCTGCGTCTACCCGCGATCTTGTGCTCTTCTGCTTCCATTGGTACACGCCCTCTC 2006
Db 415 agtgggcagcaggtctgagtaactctcatatgatgctgatttctctggagttgacagaa 474
Qy 2007 AGTGGGACGAGGCTGTGACTACTCTCATATGATGCTGATTTCTCGAGTGTACAGAA 2066
Db 475 acacctattttccccccagctcttccctggagactacttaactgaataaataactcag 534

Aug 27 08:56

US-08-231-565A-26.rgc

30

Qy 2067 ACACCTATATTCCCCAGCTCTCCCTGGGAGACTACTATTAACTGAATAAATCTCAG 2126
Db 535 agcctga 541
Qy 2127 AGCCTGA 2133
RESULT 13
LOCUS HSPMEL04 1559 bp DNA PRI 19-DEC-1995
DEFINITION Human melanocyte protein Pmel 17 gene, exons 7, 8, 9, and 10.
ACCESSION U31798
NID g1125060
KEYWORDS
SEGMENT 4 of 5
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1559)
AUTHORS Baillin, T., Lee, S.-T. and Spritz, R.A.
TITLE Genomic organization and sequence of D12553E (Pmel 17), the human
homologue of the mouse silver (si) locus
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 1559)
AUTHORS Baillin, T.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445
COMMENT Henry Mall, University of Wisconsin, Madison, WI 53706, USA
NCBI gi: 1125060
FEATURES
source
1..1559
/organism="Homo sapiens"
/chromosome="12"
/map="12q13-q14"
/note="DSEG number: D12553E"
382..496
/number=7
736..821
/number=8
932..1135
/number=9
1238..1327
/number=10

BASE COUNT 340 a 392 c 413 g 414 t
ORIGIN

Query Match 9.6%; Score 209; DB 86; Length 1559;
Best Local Similarity 100.0%; Pred. No. 1.98e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 930 ggctgcccaaggaagcctcatggagatctcatgccagggtgccagccctgccagc 989
Qy 1594 GCCTGCCCAAGGAAGCCTCATGCGAGACTCTCATGCCAGGGTGCCAGCCCTGCCAGC 1653
Db 990 ggctgtgccagcctgtgtaccagccagcagcctgccagctggctgtgccagagactaga 1049
Qy 1654 GCGTGTGCCAGCCTGTGTACCAGCCAGCCTGCCAGCTGCCAGCTTCTGCACGAGACTGA 1713
Db 1050 aggtgtggctcggggacactgcctcaatgtgtctgtgctgataccaacagcctggcag 1109
Qy 1714 AGGGTGGCTCGGGGACATACTGCCTCAATGTCTCTCTGTGCTGATACCAACAGCCTGGCAG 1773
Db 1110 tggctcagcaccagcttatcatgctggt 1138

QY 1774 TGGTCAGCACCAGCTTATCATGCTGTGT 1802

RESULT 14
ID HSMEL04 standard; DNA; PRI; 1559 BP.
AC U31798;
DT 23-DEC-1995 (Rel. 46, Created)
DT 23-DEC-1995 (Rel. 46, Last updated, Version 1)
DE Human melanocyte protein Pmel 17 gene, exons 7, 8, 9, and 10.
KW
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-1559
RA Bailin T., Lee S.T., Spritz R.A.;
RT "Genomic organization and sequence of D12S53E (Pmel 17), the human
RT homologue of the mouse silver (si) locus";
RL Unpublished.
RN [2]
RP 1-1559
RA Bailin T.;
RT
? Submitted (17-JUL-1995) to the EMBL/GenBank/DBJ databases.
RL Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of
RL Wisconsin, Madison, WI 53706, USA
CC NCBI gi: 1125060
FH Key Location/Qualifiers

FT source 1..1559
FT /organism="Homo sapiens"
FT /chromosome="12"
FT /map="12q13-q14"
FT /notes="DSEG number: D12S53E"
FT exon 382..496
FT /number=7
FT exon 736..821
FT /number=8
FT exon 932..1135
FT /number=9
FT exon 1238..1327
FT /number=10
SQ Sequence 1559 BP; 340 A; 392 C; 413 G; 414 T; 0 other;

Query Match 9.6%; Score 209; DB 10; Length 1559;
Best Local Similarity 100.0%; Pred. No. 1.98e-90;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 930 ggctgccaggagcctgcatggagatctcatcgccagggtgccagccctgccagc 989
|||||
QY 1594 GCCTGCCAAGAGAGCTGCATGAGATCTCATGCCAGGTCGCCAGCCCTGCCAGC 1653
|||||
Db 990 ggctgtgccagctgtgctaccagccagcctgccagctgttctgcacagatactga 1049
|||||
QY 1654 GGCTGTGCCAGCCTGTCTACCCAGCCAGCCCTGCCAGCTGCTTCCACCACTACTGA 1713
|||||
Db 1050 aggggtggctggggaactactcctcaatgtgtctctggctgatacacaagcctggcag 1109
|||||
QY 1714 AGGGTGGCTCGGGACATACCTGCTCAATGTCTCTGGCTGTATACCAAGCCTGGCAG 1773
|||||
Db 1110 tggtcagcaccagccttatcatgcctggt 1138
|||||
QY 1774 TGGTCAGCACCAGCTTATCATGCTGTGT 1802

RESULT 15
LOCUS HSPMEL02 673 bp DNA PRI 19-DEC-1995
DEFINITION Human melanocyte protein Pmel 17 gene, exons 2 and 3.
ACCESSION U31807
NID g1125058
KEYWORDS
SEGMENT 2 of 5
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; mitochondria; eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 673)
AUTHORS Bailin, T., Lee, S.-T. and Spritz, R.A.
TITLE Genomic organization and sequence of D12S53E (Pmel 17), the human
TITLE homologue of the mouse silver (si) locus
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 673)
AUTHORS Spritz, R.A.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445
Henry Mall, University of Wisconsin, Madison, WI 53706, USA
COMMENT NCBI gi: 1125058

FEATURES
source
1..673
/organism="Homo sapiens"
/chromosome="12"
/map="12q13-q14"
/notes="DSEG number: D12S53E"
128..241
/number=2
400..546
/number=3
BASE COUNT 186 a 142 c 179 g 166 t
ORIGIN

Query Match 6.9%; Score 150; DB 86; Length 673;
Best Local Similarity 100.0%; Pred. No. 1.46e-58;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 396 agtggtcaagtgctcccaagtcagtaagtgccctacactggtggtgcaaatgc 455
|||||
QY 224 AGGTGGTCAAGTGTCCCTCAAGGTCAGTAATGATGGGCTACACTGATTGTCGAATGC 283
|||||
Db 456 ctctctctctatgctgaacttcctccctggagccaaagattatgccagatggcaggt 515
|||||
QY 284 CTCCTTCTCTATTGGCTTGAAGTCTCCCTGGAAGCCAAAGGTTATTGCCAGATGGCAGGT 343
|||||
Db 516 tatctgggtcaacaataccatcatcaatgg 545
|||||
QY 344 TATCTGGGTCAACATATACCATCATCATGG 373
|||||

Search completed: Tue Aug 27 09:09:11 1996
Job time : 1535 secs.

Aug 27 09:00

US-08-231-565A-26.mrg

1

WAVELENGTH

(TM)

Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 09:09:30 1996; MasPar time 177.89 Seconds
868.381 Million cell updates/sec

Tabular output not generated.

Title: >US-08-231-565A-26
Description: (1-2172) from US08231565A.seq
Perfect Score: 2172
N.A. Sequence: 1 GTCCAGCGCCATTACCAATC.....AAAAAAAAAAAAAAAAAAAA 2172
Comp: CAGCTCGCGGTATGCTAG.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 91712 seqs, 35561689 bases x 2

Post-processing: Minimum Match 04
Listing first 45 summaries

Database: n-genseq23

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17

Statistics: Mean 9.903; Variance 7.604; scale 1.302

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description	Pred. No.
1	2172	100.0	2172	17	T02716	MART-1 melanoma antig	0.00e+00	
2	2101	96.7	2115	16	Q96055	Sequence encoding mel	0.00e+00	
3	1970	90.7	2131	17	T03760	Melanoma-specific imm	0.00e+00	
4	88	4.1	1047	2	Q10572	Human Natriuretic Pep	2.95e-29	
c	5	86	4.0	1047	2	Q10572	Human Natriuretic Pep	3.04e-28
6	45	2.1	458	15	Q90525	Rat SIII 15 kDa subun	2.36e-08	
c	7	44	2.0	204	1	N81164	Base substituted E.co	6.69e-08
c	8	43	2.0	350	16	Q99380	Rat allograft inflam	1.88e-07
9	43	2.0	1338	11	Q65607	Rabbit zona pellucida	1.88e-07	

Aug 27 09:00

US-08-231-565A-26.mrg

2

10	44	2.0	1710	8	Q47966	Rape acyl-ACP thioest	6.69e-08
11	43	2.0	1813	15	Q88760	Human ubiquitous nucl	1.89e-07
12	42	1.9	91	9	Q51746	Oligonucleotide probe	5.27e-07
13	42	1.9	91	9	Q51746	Oligonucleotide probe	5.27e-07
14	41	1.9	699	10	Q58663	Human superoxide dis	1.46e-06
15	41	1.9	756	12	Q73729	Rape abscission/dehis	1.46e-06
16	42	1.9	861	17	Q99552	Thrombopoietin coding	5.27e-07
17	41	1.9	971	13	Q73760	RP-8 Programmed cell	1.46e-06
18	41	1.9	1325	5	Q31873	Cyclin D1 gene.	1.46e-06
19	41	1.9	1325	9	Q53197	Human cyclin D1.	1.46e-06
20	41	1.9	1394	1	Q05879	Placenta-specific pro	1.46e-06
21	41	1.9	1972	16	Q95309	mPD-1 cDNA.	1.46e-06
22	41	1.9	1973	10	Q55770	Murine programmed cel	1.46e-06
23	41	1.9	2377	3	Q20360	Human pro-urokinase c	1.46e-06
24	42	1.9	2381	12	Q68910	Human c-myc far upstr	5.27e-07
25	42	1.9	2384	12	Q68909	Human c-myc far upstr	5.27e-07
26	41	1.9	2427	1	Q04107	Human pro-urokinase c	1.46e-06
27	41	1.9	2696	2	N60889	Plasmid pAU157 sequen	1.46e-06
28	41	1.9	2696	2	N60909	Plasmid pAU157 insert	1.46e-06
29	41	1.9	2697	2	N60864	Sequence of plasmid p	1.46e-06
30	41	1.9	3138	2	Q11712	Shuttle vector pMW16	1.46e-06
31	42	1.9	3504	1	Q05304	Clone Imd2 encoding i	5.27e-07
32	40	1.8	635	13	Q75445	Murine glycosylation	4.04e-06
33	40	1.8	667	11	Q61575	Hybrid human/pig SOD.	4.04e-06
34	40	1.8	688	11	Q61572	Human SOD.	4.04e-06
35	40	1.8	700	3	Q22352	P14(T) allergen clone	4.04e-06
36	40	1.8	887	1	Q05160	Fragment of plasmid p	4.04e-06
37	40	1.8	1393	6	Q38027	CAD cDNA isolated fro	4.04e-06
38	40	1.8	1576	8	Q49944	Human anti-IBS heavy	4.04e-06
39	40	1.8	2339	11	Q63892	Tomato acid invertase	4.04e-06
40	40	1.8	2431	5	Q31546	Human heregulin-beta2	4.04e-06
41	40	1.8	2445	7	Q48393	Full-length foetal hu	4.04e-06
42	40	1.8	3207	3	Q14263	p-meta-1 metastasis-s	4.04e-06
43	40	1.8	3312	2	N71034	pMx34 cDNA insert.	4.04e-06
44	40	1.8	3628	2	N60848	Plasmid sequence enco	4.04e-06
45	40	1.8	4589	1	Q06068	p150.95 alpha subunit	4.04e-06

ALIGNMENTS

RESULT	1
ID	T02716 standard; cDNA; 2172 BP.
AC	T02716;
DT	20-APR-1996 (first entry)
DE	MART-1 melanoma antigen cDNA25.
KW	cDNA25; MART-1; melanoma antigen recognised by T-cells;
KW	gp100 antigen derivative; melanoma; metastatic melanoma;
KW	tumour-associated antigen; immunogen; diagnosis; prognosis;
KW	prophylaxis; therapy; vaccine; ds.
OS	Mammalian.
FH	Key
FT	Location/Qualifiers
FT	CDS
FT	/*tag= a
FT	/*note= "cDNA25 melanoma antigen"
PN	W09529193-A2.
PD	02-NOV-1995.
PF	21-APR-1995; U05063.
PR	22-APR-1994; US-231565.
PR	05-APR-1995; US-417174.
PA	(USSH) US SEC DEPT HEALTH.
PI	Kawakami Y, Rosenberg SA;
DR	WPI; 95-382963/49.
DR	P-PSDB; R84854.
PT	DNA encoding melanoma antigens recognised by T-lymphocytes - also

PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Disclosure; Fig 4A-4B; 184pp; English.
CC The nucleic acid encodes cDNA25, a melanoma antigen (MART-1)
CC which is recognized by T-lymphocytes. cDNA25 is a derivative of
CC the melanocyte-melanoma-specific antigen gp100 (see R84855).
CC Antigen cDNA25 is a source of immunogenic peptides (see R84199)
CC which are optionally modified (see R84200-R84211) and used in
CC medicaments, especially vaccines, for the treatment or prevention
CC (by immunization) of melanoma. Antibodies against cDNA25 and its
CC immunogenic peptides may be used in the detection and isolation
CC of the antigen from a sample, the detection of which is indicative
CC of a disease state (melanoma or metastatic melanoma).
SQ Sequence 2172 BP; 512 A; 594 G; 578 C; 488 T;

Query Match 100.0%; Score 2172; DB 17; Length 2172;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 gtcagcgccattaccatcgaccggggaagacacacatggatctgtgtcctaaagat 60
|||||
Qy 1 CTCGACGGCCATTACCAATCGCCGCGGGAAGACACATGGATCTGGTCTAAAGAT 60
|||||
Db 61 gctcttctcatttgctgatagtgcttgcgtgtgggtggtcctaaagatcacca 120
|||||
Qy 61 GCCTTCCTTATTGGCTGTGATAGTGGCTTTGCTGGCTGGGGGCTCAAAAGTAGTACCA 120
|||||
Db 121 gaaacacagagctgcttggctcctcaagcaactcagaacaaagcctggaacagcagc 180
|||||
Qy 121 GAAACAGGACTGGCTGGTGTCTCAGGCAACTCAGAAACCAAGCCTGGAACAGGCAGC 180
|||||
Db 181 tctatccagagtgacacagaagccagagactgactgctggagagtggtcagaagtccc 240
|||||
Qy 181 TGTATCCAGAGTGGACAGAGCCAGACTTGACTGCTGGAGAGTGGTCAAGTGTCCC 240
|||||
Db 241 tcaaggtcagtaagtggcctacactgattgtgcaaatgctcctctctattgct 300
|||||
Qy 241 TCAGGTGCAATGATGNGGCCCTACACTGATTTGGTGCAATGGCTCCTCTCTATTGGCT 300
|||||
Db 301 tgaacttcctggagaccaaagattggccagatggcaggttactgggtcaacaata 360
|||||
Qy 301 TGAACCTTCCCTGGAAGCCAAAGGATTTGCCAGATGGCAGGTTATCTGGGTCAACAATA 360
|||||
Db 361 ccatcatcaatgggagcagtgctgggagagcagccagtgatccccaggaactgacg 420
|||||
Qy 361 CCATCATCAATGGGAGCCAGCTGTGGGAGGACAGCCAGTGTATCCCAGGAACTGAGG 420
|||||
Db 421 atgctgcattctccctgatgtggacctggccatctggctcttggctcagaagaaa 480
|||||
Qy 421 ATGCTCGCATCTCTCCGATGTGTGGACCTTGGCCACTCTGGCTCTTGGTCTCAGAGAAA 480
|||||
Db 481 gctttgttatgtctggaagacctgggccaactactggcaatttctagggggccagtg 540
|||||
Qy 481 GCTTTGTTTATGTCTGGAGAACCTGGGGCCCAATCTTGGCAATTTCTAGGGGGCCAGTGT 540
|||||
Db 541 ctgggctgaacattggagacgaggggcaatgctgggcacacacacacatggaagtga 600
|||||
Qy 541 CTGGGCTGACCAATTGGAGAGCGAGGGCAATGCTGGGACACACACCAATGGAAGTCACTG 600
|||||
Db 601 tctaccatcgccggggtatccggagctatgtgcctcttgctcattccagctcagcctca 660
|||||
Qy 601 TCTACATCCGCGGGGATCCCGGAGTATGTGCTCTTGTCTTCCATTCCAGCTCAGCCTTCA 660
|||||
Db 661 ccattactgaccaggtgctcttctccgtgagcgtgtcccagttgcgggcttgatggag 720

Qy 661 CCATTACTGCACAGGTGCTTTCTCGTGACGGGTGCCAGTTGCGGGGCTTGGATGGAG 720
|||||
Db 721 ggaacaagcacttccctgagaatcagcctctgaccttggccctccagctccatgacccca 780
|||||
Qy 721 GGAACAAGCACTTCTCTGAGAAATCAGCCTCTGACCTTTGGCCCTCCAGCTCCATGACCCCA 780
|||||
Db 781 gtgtctatctggctgaagctgacctctctacacctgggaactttggadacagtagtggaa 840
|||||
Qy 781 GTGGTATCTGGCTGAACCTGACCTCTCTACACCTGGGACTTTGGAGACAGTAGTGGAA 840
|||||
Db 841 cctgatctctcgggcacttgggtcactcatacttacctggagcctggcccaagtcaetg 900
|||||
Qy 841 CCTGATCTCTGGGCACTTGTGGTCACTATATTACTTACCTGGAGGCTGGCCCACTCACTG 900
|||||
Db 901 cccagtggtcctgagggctgacctctcctcactcctgtggctcctccccaggttccag 960
|||||
Qy 901 CCCAGTGTCTCTGACGGCTGCCATTCCTCTCACCCTCTCTGTGGCTCTCCCAAGTTCCAG 960
|||||
Db 961 gcaacacagatgggcacagggccaactgcagaggccctaacacacacagctggccaagtgc 1020
|||||
Qy 961 GCACACAGATGGGCACAGGCCCAACTGCGAGAGGCCCTTAACACCAAGCTGGCCAAAGTGC 1020
|||||
Db 1021 ctactacagagtggtggtaactacacctggtcagcgcccaactgcagagccctctggaa 1080
|||||
Qy 1021 CTACTACAGAGTGTGGGTACTACACCTGCTGAGCGCCCACTGCGAGAGCCCTCTGGAA 1080
|||||
Db 1081 ccacatctgtcaggtgccaacccactgaagtcatagcactgcacctgtcgagatgccaa 1140
|||||
Qy 1081 CCACATCTGTCTGAGGTGCCAACCACTGAAGTCAATAAGCACTGCACTGTGCAGATGCCAA 1140
|||||
Db 1141 ctgcagagacacaggtatgacacctgagaagtgccagtttcagaggtcatgggtacca 1200
|||||
Qy 1141 CTGCAGACACAGGTATGACACCTGAGAAGGTGCCAGTTTCAGAGGTCTATGGGTACCA 1200
|||||
Db 1201 cactggcagagatgtcaactccagaggctacaggtatgacacctgcagagtgatcaattg 1260
|||||
Qy 1201 CACTGGCAGAGATGTCAACTCCAGAGGCTACAGGTATGACACCTGCGAGAGGTATCAATTG 1260
|||||
Db 1261 tgggtctttctggaaccacagctgcacaggttaacaactacagagtggtggagaccacag 1320
|||||
Qy 1261 TGGTCTCTTCTGGAACCAACAGCTGCCAGGTTAAACACTACAGAGTGGGTGGAGACCACAG 1320
|||||
Db 1321 ctgagagagctaacatccctgagcctgaaggtccagatgcccagctcaatcatgtctacgg 1380
|||||
Qy 1321 CTAGAGAGCTACCTATCCCTGAGCCTGAAGTCCAGATGCCAGCTCAATCATGTCTACGG 1380
|||||
Db 1381 aaagtattacaggttccctgggccccctgctggatgggtacagccaccttaaggctggtga 1440
|||||
Qy 1381 AAGTATTACAGTTCCCTGGGCCCCCTGCTGCTGATGCTAGACGCCACCTTAAGGCTGGTGA 1440
|||||
Db 1441 agagacaagtccccctggattgttctgtatcgatagttgctcttttccgtcacctgg 1500
|||||
Qy 1441 AGAGACAAGTCCCCCTGGATTTGTTCTGTATCGATATGGTTCCTTTTCCCTCACCTGG 1500
|||||
Db 1501 acattgtccaggggtattgaaagtgcagagatccctcagagctgtgcgtccggtgaggggg 1560
|||||
Qy 1501 ACATTGTCCAGGGTATTGAAAGTGCCTGAGATCCTGCAGGGCTGTGCCGCTCGGTCAGGGGG 1560
|||||
Db 1561 atgcatcttgagctgactgtcctgcaagggcggtgctcccaaggaagcctgcagtggaga 1620
|||||
Qy 1561 ATGCATTTGAGCTGACTGTCTCTGCCAAGGGGGCTGCCAAAGGAAGCCTGCTGATGGAGA 1620
|||||
Db 1621 tetcatcgacaggtgcaagccccctgccacagcgctgtgccagctgtgctaccacagcc 1680

```

Qy 1621 TCTCATGCCAGGTCAGCCGCTGCGCAGCGGCTGCGCAGGCTGTGCTACCGGCC 1680
Db 1681 cagcctgccagctggttctgaccagatactgaaggggtgctcggggacatactgctca 1740
Qy 1681 CAGCCTGCCAGCTGCTTCTGCACAGATACTGAAGGGTGGCTCGGGGACATACCTGCTCA 1740
Db 1741 atgtctctggctgatacaacagcctggcagtggtgagcaccagcttatcatgctg 1800
Qy 1741 ATGTGCTCTGGCTGATACCAACAGCCTGGCAGTGTGAGCAGCCAGCTTATCATGCTG 1800
Db 1801 gtcaagaagcagccttggcaggttcgctgctgctggtggtggtggtggtggtggtg 1860
Qy 1801 GTCAAGAAGCAGCCCTTGGCAGGTTGGCCTGATCGTGGCATCTTGGTGTGTTGATGG 1860
Db 1861 ctgtggtcctgcatctctgatataagcgcagacttatgaagaagacttctcgttac 1920
Qy 1861 CTGTGCTCTGCTCATCTCTGATATATAGCGCAGACTTATCAAGCAAGACTTCTCGGTAC 1920
Db 1921 cccagttgccacatagcagcagctgctgctgctaccccgcatcttctgtctgttc 1980
Qy 1921 CCCAGTTGCCACATAGCAGCAGTCACTGGCTGGCTGCTACCCCGCATCTTCTGCTTGT 1980
Db 1981 ccattggtgagaaagccctcctcagtgagcagcagctgctgactctcatatgatg 2040
Qy 1981 CCATTGTTGAGAACAGCCCTCTCAGTGGCAGCAGGCTCAGTACTCTCATATGATG 2040
Db 2041 ctgtgatttctcctgagtgagaaacactatatttccccagcttctcctggagac 2100
Qy 2041 CTGTGATTTCTCTGAGTTGACAGAAACACTATATTTCCCCAGTCTTCCCTGGCAGAC 2100
Db 2101 tactattactgaaataaataactcagagcctgaaaaataaaaaaataaaaaa 2160
Qy 2101 TACTATTAACTGAATAAATACTCAGACGCTGAATAAATAAATAAATAAATAA 2160
Db 2161 aaaaaaataaata 2172
Qy 2161 AAAAAAATAAATA 2172

RESULT 2
ID Q96055 standard; cDNA to mRNA; 2115 BP.
AC Q96055;
DE Sequence encoding melanoma associated antigen gp100.
KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
KW identification; tumour; gp100; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS
FT /*tag= a
FT /product= Melanoma associated antigen gp100.
FT misc_signal 1..81
FT /*tag= b
FT misc_feature 1792..1870
FT /*tag= c
FT /label= Transmembrane domain.
FT misc_binding 262..264
FT /*tag= d
FT /bound_moeity= Carbohydrate.
FT misc_binding 337..339
FT /*tag= e
FT misc_binding 352..354
FT /*tag= f

```

```

FT misc_binding 982..984
FT /*tag= g
FT misc_binding 1723..1725
FT /*tag= h
PN EP-668350-A1.
PD 23-AUG-1995.
PF 14-FEB-1995; 200348.
PR 16-FEB-1994; EP-200337.
PR 21-DEC-1994; EP-203709.
PA (ALKU ) AKZO NOBEL NV.
PI Adema GU, Flgdon CG;
DR WPI; 95-284790/38.
P-PSDB; R78646.
PT Melanoma associated antigen gp100 - used in vaccines and for the
PS Claim 2; Page 19-22; 40pp; English.
CC Immunogenic peptides derived from the melanoma associated antigen
CC may be used in the production of vaccines. Nucleotide sequences
CC encoding the immunogenic peptides may be used as primers and probes
CC in the detection of melanoma cells. Tumour infiltrating lymphocytes
CC capable of binding to the melanoma associated antigen can be
CC cultured ex vivo and returned to melanoma particles, and when
CC radiolabelled, they may be used to identify tumour deposits.
SQ Sequence 2115 BP; 469 A; 587 C; 575 G; 484 T;

Query Match 96.7%; Score 2101; DB 16; Length 2115;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 11 ggaagaacacaatggatgctggtctaaaaagatgctcttcttcttctgctgtagtggtg 70
Qy 28 GGAAGAACAATGGATCTGGTCTAANAAGATGCGCTTCTTCATTTGGCTGTGATAGGTG 87
Db 71 ctttctgctgtggggctacaaaagtaccagaaaccagagactgcttgggtgtctcaaa 130
Qy 88 CTTTCTGCTGTGGGGCTACAAAAGTACCAGAAACCAAGGACTGGCTTGGTGTCTCAA 147
Db 131 ggcactcagaacaaaagcctggaacaggcagctgtatccagagtggacagaagccaca 190
Qy 148 GGCMACTCAGAACCAAGCCTGGAAACAGGCAGCTGTATCCAGAGTGGACAGAGCCACA 207
Db 191 gacttgactgtggagaggtgggtcaagtgtccctcaaggtcagtaagtatgatggcctcac 250
Qy 208 GACTTGACTCTGGAGAGGTGTCAGGTGTCCTCAAGGTCAGTATGATGGGCTTACAC 267
Db 251 tgattggtgcaaatgctcctctctctattgcttgaacttccctggaagcacaaggtat 310
Qy 268 TGATTGGTGCMAATGCTCCTCTCTATTTGCTTTGAACTTCCCTGGAGAGCCCAAGGTAT 327
Db 311 tgccagatgggaggttatctctgggtcaacaataccatcatcaatggagccaggtgtggg 370
Qy 328 TCCAGATGGCAGGTTATCTGGGTCAACAATATCATCAATGGCAGCCAGGCTGTGGG 387
Db 371 gaggacagccagtgatccccaggaaaactgacatgcctgcattctcctcctgatggtagc 430
Qy 388 GAGGACAGCCAGTGTATCCCGACGAAACTGACGATGCCTGCATCTTCCCTCATGTTGGAC 447
Db 431 cttgccatctggctcttggtctcagaagaagagcttggttttatgtctggaagacctggg 490
Qy 448 CTTGCCCATCTGGCTCTTGGTCTCAGAGAGAGAGCTTTGTTTATGTCGTGAAGACCTGGG 507
Db 491 gccaatactggcaagttctagggggccagctgtctgggtgagcattgggacagggcaggg 550
Qy 508 GCCAATACTGGCAATTTCTAGGGGGCCCGACGTGTCTGGGCTGAGCATTTGGGACAGCAGG 567

```

Aug 27 09:00

US-08-231-565A-26.mg

D	b	551		caatgctgggcacacacaccattgaagtgactgtacctcgcgggatcccgagct	610
		552			
		553			
Q	y	558		CAATGCTGGGCACACACACCATTGGAGTGACTGTCTACCATCGCCGGGATCCGGGAGCT	627
		559			
		560			
D	b	611		atgtgctcttgcctcatccagctcagccttcaccattactgaccagggtgcctttctccq	670
		612			
Q	y	628		ATGTCGCTTCTGCTCATTTCCAGCTCAGCGCTTCACATTACTGACAGGTGCTTTCTCGG	687
		629			
D	b	671		ttagcgtgtcccagttgcggggccttggatgaagggaacaacgacctcctgagaatacag	730
		672			
Q	y	688		TGAGGCTGTCCCAGTTGGTGGGGSCCTTGGATGGAGGGGAACAAGCACTTCTCTGAGAATAACG	747
		689			
D	b	731		ctctgaaccttgtccctccagctccatgaacccagtgcctatctgcgtgaagctgaacct	790
		732			
Q	y	748		CTCTGACCTTTGCCCTTCCAGCTCCATGACCACGTGGCTATCTGGCTGAAGCTGACCTCT	807
		749			
D	b	791		cctaacacctgggactttggagacagtaagtgaacctgatctctcgggcactttggtgcca	850
		792			
Q	y	808		CCTACACCTGGGACTTTGGAGACAGTAGTGGNAACCTGATCTCTCGGGCACTTGTGTCCA	867
		809			
D	b	851		ctcatactactgcgagcctggcccagtcactgccaggtggtcctgcaggctgccattc	910
		852			
Q	y	868		CTCATTACTTACTGTGAGCGCTGGCCCCAGTCACTGCCCAGGTGGTCTCTGACAGGCTGCCATTC	927
		869			
D	b	911		ctctcaectcctgtggctcctccccagtttcaggaccacacagatggcacaggcccaactg	970
		912			
Q	y	928		CTCTCACCTCTCTGTGGCTCTCTCCCAAGTTCCAGGCAACACAGATGGGCAAGGCCAAMCTG	987
		929			
D	b	971		cadaggccctaaacaccacagctggccaaagtgcactacagaagttgtgggtactaac	1030
		972			
Q	y	988		CAGAGGCCCTTAACACCAACAGCTGGCCAAAGTGCTACTACAGAAAGTTGTGGGTACTAAC	1047
		989			
D	b	1031		ctggtcagggcaccaactgcagagccctctgaaaccaatctgtgcagggtgccacaacactg	1090
		1032			
Q	y	1048		CTGGTCAGGGGCCAACTGCAGAGCCCTCTGGAACCAACATCTGTCAAGTGGCAACCACTG	1107
		1049			
D	b	1091		aagtcataaagcactgcacctgtgcagatgccaaactgcagagagacaggtatgcacctg	1150
		1092			
Q	y	1108		AAGTCATAAGCACTGCACCTGTGCAGATGCCAACTGCAGAGACACAGGTATGCACCTG	1167
		1109			
D	b	1151		agaaggtgccagtttcagaggtcatgggtaccacctggcagagatgtcaactccagagg	1210
		1152			
Q	y	1168		AGAAAGTGCCANGTTTCAGAGGT CATGGGTATCCACACTGGCAGAGATGTCAACTCCAGAG	1227
		1169			
D	b	1211		ctcaggttatgacacctgcagaggtatcaattgtggtcttcttggaaccaacagctgcac	1270
		1212			
Q	y	1228		CTACAGGTATGCACCTGCCAGAGGTATCAATTGTGTGCTTTCTGGNACACACACTGCAC	1287
		1229			
D	b	1271		aggtaaacactacagatgggtggagacacagctagagactacctatccctgagcctg	1330
		1272			
Q	y	1288		AGGTAAACAACCTACAGAGTGGGTGGAACCAACAGCTAGACAGCTACTCTATCCCTGAGCGCTG	1347
		1289			
D	b	1331		aaggtccagatgccagctcaatcatgtctacggaagattatacaggttccctgggcccc	1390
		1332			
Q	y	1348		AAGTCCAGATGCCAGCTCAATCATGTCTACGGNAAAGTATTACAGGTTCCCTGGGCCCCC	1407
		1349			
D	b	1391		tgctggatggtacagccacctaaagctgggtgaagagacaagtcacctggatgtgttcc	1450
		1392			
Q	y	1408		TGCTGGATGTGTACAGCCACCTTAAGCTGTGTTAAGAGACACAGTCCCCTTGATTTGTGTTCT	1467
		1409			
D	b	1451		tgtatcatgatgtctcttttccgctcacctggacattgcagggtattataaagtgccg	1510
		1452			
Q	y	1468		TGTTATGCAATGTTCTCTTTTCCGTCACCCCTGCACATCTCTCCAGGGTATTGAAAAGTGGCG	1527
		1469			

Db	1511	agatcctgcaggctgtgcgcgtccggtgcaggggatgcatttcgagctgactgtctcctgcc	1517
Qy	1528	AGATCCTGCAGGCTGTGCCGTGCGGTGAGGGGATGCATTTTCACTCACTGTGTCTCGCC	1587
Db	1571	aagcgggctgtcccaaggaaagcctgcattggagatctcatgccagggtgccagccccctg	1630
Qy	1588	AAGCGGGCTGCCCAAGGAGCCTGCATGCAGATCTCATGCCAGGGTGCCAGGCCCGCTG	1647
Db	1631	cccagcgctgtccagcctgtctaccacgccagcctgccagctggctgcaccaa	1690
Qy	1648	CCACGGGGCTGTGCCAGCCTGTGCTACCAACCCACGCTGCCAGCTGGTTCTGCACCA	1707
Db	1691	tactgaagggtggctcggggacatcctgcctcaatgtctctggctgataccaacagcc	1750
Qy	1708	TACTGAAGGGTGGCTCGGGACATATCGCTCAATGTGTCTGGCTGATACCAACAGCC	1767
Db	1751	tggcagtggtcagcaccagcttatcagcctggctcaagaagcagcccttgggcaggttc	1810
Qy	1768	TGGCAGTGGTCAGCACCCAGCTTATCATGCTGGTCAAGAAGCAGGCCCTTGGCAGGTT	1827
Db	1811	cgtgtagctggggcatcttgcctggtgtgatggcttggctccttgcactctgatata	1870
Qy	1828	CGGTGATGCTGGGCATCTTGTGGTCTTTGATGGCTGTGGTCTTGCATCTCTGATATA	1887
Db	1871	ggcgcagactatgaagcaagactctcgtaccagcttgccacatagcagcagtcact	1930
Qy	1888	GGCGCAGACTTATGAAGCAAGACTTCTCCGCTACCCCACTGCCACATCAGGAGTCACT	1947
Db	1931	ggctgcgtctacccegcactcttgccttgcatttggtagaataagccccctcctca	1990
Qy	1948	GGCTGGCTCTACCCCGCATCTTGTCTTTCCTGCGATGGGTGAAACAGCCCCCTCCTCA	2007
Db	1991	gtgggcagcaggtctgagtactctcatalgatgctgattttcctggagtgacagaa	2050
Qy	2008	GTGGGCACGAGGTCTCAGTACTCTCATATGATGCTGTGATTTTCTGGAGTTGACAGAA	2067
Db	2051	caactatatttccccagctcttccctgggagactactattaactgaataaataactcaga	2110
Qy	2068	CACCTATATTTCGCCCGAGCTTCCCTGGGAGACTACTATTAACTGAAATAAATACTCAGA	2127
Db	2111	gcctg	2115
Qy	2128	gcctg	2132

RESULT 3

ID T03760 standard: DNA: 2131 BP.

ID T03760
AC T03760:

DT 25-MAR-1996 (first entry)

DE Melanoma-specific immunogen, pMEL17, 25-ETH-1990 (ATCC entry)

KW Melanoma: immunogen: epitope; homologue: vaccine; immunotherapy;

KW cytotoxic T cell: lymphocyte; HLA-A2: ss.

OS Homo sapiens.

PN W09522561-A2:

PD 24-AUG-1995.

PF 16--FEB-1995; U01991.

PR 16-FEB-1994; US-197399.

PR 29-APR-1994; US-234784.

PA (UYVI-) UNIV VIRGINIA PATENT FOUND.

PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;

DR WPI; 95-302688/39.

PT Melanoma-specific immunogen comprises epitope(s) homologous with

PT pMel.17 -- are highly potent stimulators of HLA-A2+CTL's useful in

PT adoptive immuno-therapy
PS Disclosure; Page 19-20; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 (t03760) comprises
CC one or more CTL (cytotoxic T lymphocyte) epitopes from the group
CC R82098-R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines. 491
SQ Sequence 2131 BP; 474 A; 589 C; 577 G; 491 T;

Query Match 90.7%; Score 1970; DB 17; Length 2131;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 2103; Conservative 0; Mismatches 7; Indels 21; Gaps 1;

Db	1	ggaagaacacatggtatgctgctaaaaagatgccttctcatttggctgtagtggtg 60
Qy	28	GGAGACACATGATCGTGGTCTAAAGATGCCCTTCATTTGGCTGTGATAGGTG 87
Db	61	ctttgctggctgggggtacaaaagtcaccagaaccaggactggttggctgtctcaa 120
Qy	88	CTTTGCTGGCTGGGGGCTACAAAAGTACCCAGAACAGAGACTGGCTGGTGGTCTCAA 147
Db	121	ggcaactcagaacaaagcctggaacagcagctgtatccagatggacagaagccaga 180
Qy	148	GCCAACTCAAGAACCAAGGCTGGAACAGGAGCTGTATCCAGAGTGGACAGAGCCGAGA 207
Db	181	gacttgactgctggagagtggtcgaagtgcctcaagtcagtcgaatgatggctcaac 240
Qy	208	GACTTGACTGCTGGAGAGTGGTCAAGTGTCCCTCAAGGTCAGTAATGATGGGCTACAC 267
Db	241	tgattggtgcaaatgcctcctctctattgcttgaacttccttggagagccaaaagtat 300
Qy	268	TGATTTGGTCAAAATGCCCTCTCTCTATTGGCTTGAATTCCTCGGAGCCAAAAGGAT 327
Db	301	tgccagatggcaggttatctgggtcaacaataccatcatcaatggagccaggtgtggg 360
Qy	328	TCCAGATGGGAGGTATTCTGGGTCAACATATACCATCATCATGGAGCCAGCTGGG 387
Db	361	gaggaagcagctgtatccccaggaactgacatgcttgcattcttcctgattggtgac 420
Qy	388	GAGGACAGCCAGTGTATCCCCAGGAACTGACGATGCCCTGCATCTTCCTCATGSGTGAC 447
Db	421	cttgcccatctggtcttggctcagaagagaagctttgtttatgtctcgaagacctggg 480
Qy	448	CTTGCCCATCTGGCTCTTGGTCTCAGAGAGAAAGTTGTTTATGTCTGGAAGACCTGG 507
Db	481	gccaactctggcaagtctcaggggcccagctgtctgggtgagcattggagcaggcaggg 540
Qy	508	GCCAACTACTGGCAATTTCTAGGGGGCCCACTGCTGCGGCTGACCATGGGACAGGCAGG 567
Db	541	caatgctgggcac 600
Qy	568	CAATGCTGGGCGACACACACATGGAAGTGAATGTCTTACCATCGCCGGGGATCCCGAGCT 627
Db	601	atgtgcctcttgcctcattccagctcagccttcacattactgacacagtgcccttctccg 660
Qy	628	ATGTGCCTCTTGTCTATTCAGCTCAGCTTTCACATTACTGACAGGTGCTTTCTCCG 687
Db	661	tgagcgtgtcccagttgcccgtcttgatogaggggaacaagcactctctgagaatcagc 720
Qy	688	TGAGCGGTGTCCAGTTGGGGGCTTGGATGGAGGGGAACAGCACTTCTCTGAGAAATCAGC 747

Db	721	ctctgacactttgccctccagctccatgaccctagtggctctatctggctgaagctgacct 780
Qy	748	CTCTGACCTTTGCCCTCCAGCTCCATGACCCAGTGGGTATCTGCTGAAGCTCACCTCT 807
Db	781	ctcacactgggactttggagacagtagtgaacctgatctctcctcctcctcctcctcctc 840
Qy	808	CCTACACCTGGGACTTTGGAGACAGTAGTGAACCTCTGATCTCTCGGGCACTTGTGCTCA 867
Db	841	ctcataactacctggagcctggccagtcactgccaggtggctcctcctcctcctcctcctc 900
Qy	868	CTCATCTTACCTTGGAGCCTGGCCAGCTACTCTCCAGGTGGTCTCGCAGGCTGCCATTC 927
Db	901	ctctcacctcctgtggtcctctcccaagttccaggtccacacagatgggcacagggccaagt 960
Qy	928	CTCTCATCTCTGTGGCTTCTCCCAAGTTCAGGCCACACAGATGGGCACAGGCCCAACTG 987
Db	961	cagagggccctaaacacacacagctggccaaagtgcctactacaagaagttgtgggtactaac 1020
Qy	988	CAGAGGCCCTTAACACACAGCTGGCCAAAGTGCCTACTACAGAAAGTTGTGGGTACTACAC 1047
Db	1021	ctggtcagggcccaactgcagagcctcctggaaccacatctgtgcaagtgccaaccactg 1080
Qy	1048	CTGGTCAGGGGCCCACTGCAGAGCCCTCTGGAACACACATCTGTGAGGTGGCAACCACTG 1107
Db	1081	aagtataagcactgcactgtgcagatgccaactgcagagacacacaggtatgacacctg 1140
Qy	1108	AGTCATAAGCACTGCCACTGTGCAGATGGCCAACTGCAGAGACACAGGTATGACACCTG 1167
Db	1141	agaaggtgccagtttcagaggtcactgggtaccacactggcagagatgtcaactccagagg 1200
Qy	1168	AGAAGGTGCCAGTTTCAGAGGTCATGGGTACCACTGCGCAGAGATGTCAACTCCAGAGG 1227
Db	1201	ctcacagtatgacactgcagaggtatcaattgtgggtccttctcctcctcctcctcctcct 1260
Qy	1228	CTACAGGTATGACACTGCAGAGGTATCAATTTGTGTGCTTCTTCTGCAACACACAGCTGCAC 1287
Db	1261	aggttaacaactacagatgggtggagaccacagctagagagtcacctatccctgagcctg 1320
Qy	1288	AGGTAAACACTACAGTGGGTGGAGCACACACTAGAGAGCTACCTATCCCTGAGCCTG 1347
Db	1321	aaggtccagatgccagctcaatcatgtctacggaaagtattacaggttccctgggcccc 1380
Qy	1348	AAGGTCAGATGCCAGCTCAATCATGTCTACGGAAAGTATTACAGGTTCCCTGGGCCCCC 1407
Db	1381	tgctggatgtacagcaaccttaaggtcgtgagagacaagtcacctggattgtgttc 1440
Qy	1408	TGCTGGATGTGTACAGCCACTTAAGGCTGGTGAAGACAAAGTCCGCCCTGCAATGTGTTC 1467
Db	1441	tgctagatagttcctcttccgtcaacctggacatgtccaggggtattgaaagtgcg 1500
Qy	1468	TGTAATGATATGTTCTCTTTCCGTCACTCGACATTTGCCAGGGTATTGAAAGTGCGG 1527
Db	1501	agatcctgcagcgtgtgcctccggtgaggggagtcattttagctgactgtgcctgcc 1560
Qy	1528	AGATCCTGCAGGGCTGTGCCGCTCGCGGTGAGGGGATGCAATTTGAGCTGACTGTCTCTGCC 1587
Db	1561	aagcgggggtgcccaaggaagcctgcattgagatctcatgccaggggtgccagccccctg 1620
Qy	1588	AAGCGGGGCTGCCAAAGGAAGCCGTGCATGAGATCTCATGCCCAAGGTCGCCAGCCCCCTG 1647
Db	1621	cccaggggtgtgccagcctgtgctacctaccagccagcctgccagctggttctgcaccaga 1680
Qy	1648	CCCAGGGGCTGCCAGCCTGTCTGTCTACCCAGCCAGCCTGCCAGCTGGTTCCTGCACAGA 1707

D	b		93	hyrrmrbnvrydynt-adaawycyrrsvkydcynachhdhhyybbbyvvnhnmc	151
C	p	:	:	: : : : : :: :: :: : : :: : : :	:
G	p	600	CAGTCACTTCACATGTGTTGGCCAGCATTCGCCTCGCCTAATGCTCAGCCCAG	541	
D	b	152	ncccbnnhvchnvbhnbnrwnayvrhdarrddvhcvc	188	
C	p	:	: : : : : : : : : : : :: :	:	
G	p	540	ACACTGGGCCGCCCTAGAATAATTGCCAGTAATTGCCCCCA	504	
R	E	S	U	L	T
RESULT 8					
ID	Q99380 standard; cDNA; 350 BP.				
IC	Q99380;				
DT	08-FEB-1996 (first entry)				
DE	Rat allograft inflammatory factor 2 (AIF-2) cDNA fragment.				
KW	AIF-2; allograft inflammatory factor 2; transplant rejection;				
KW	inhibitor; immunogenic; detection; diagnosis; ds.				
OS	Rattus sp.				
PN	M09517506-A1.				
PD	29-JUN-1995.				
PF	21-DEC-1994; U14724.				
PR	21-DEC-1993; US-171385.				
PA	(HARD) HARVARD COLLEGE.				
PI	Russell ME, Utane U;				
DR	WPJ; 95-240668/31.				
PT	DNA encoding allograft rejection factors and immunogenic fragments -				
PT	useful for identifying transplant rejection inhibitors				
PS	Claim 11; Page 92; 138pp; English.				
CC	Q99372-Q99393 are rat allograft inflammatory factor 2 (AIF-2) cDNA				
CC	fragments. The AIF-2 gene is a differentially expressed allograft gene				
CC	which is expressed in allograft tissue during transplant rejection.				
CC	Identification of the rat AIF-2 product (R60521) or transcript				
CC	indicates that allograft rejection is taking place. The rat AIF-2 gene,				
CC	fragments of this gene and/or the gene product are therefore useful in				
CC	the diagnosis of transplant rejection.				
CC	The diagnostic methods used allow rejection (vascular inflammation)				
CC	to be detected at an early stage and require only a small amount of				
CC	biopsy material.				
CC	(Sequence given in sequence ID listing in the specification gives				
CC	nucleotide 111 as C)				
SQ	Sequence 350 BP; 121 A; 37 C; 63 G; 129 T;				
Query Match 2.0%; Score 43; DB 16; Length 350;					
Best Local Similarity 84.1%; Pred.No.1.88e-07;					
Matches 53; Conservative 0; Mismatches 10; Indels 0; Gaps 0					
D	b	2	tcttaccacagatattttt	61	
C	p	:	: : :~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~	:	
G	p	2172	TTTCAGGCTCGAGTATTATT	2113	
D	b	62	tag 64		
C	p	2112	CAG 2110		
R	E	S	U	L	T
RESULT 9					
ID	Q65607 standard; cDNA; 1338 BP.				
IC	Q65607;				
DT	31-JAN-1995 (first entry)				
DE	Rabbit zona pellucida ZPC coding sequence.				
KW	Rabbit; lapine; zona pellucida; ZPC; immunocontraception; ds.				
OS	Oryctolagus cuniculus.				
FH	Key Location/Qualifiers				
CP	DS 17..1264				

```

FT /*tag= a
FT /product= ZPC
PN W09411019-A.
PD 26-MAY-1994.
PF 06-NOV-1993; U10851.
PR 09-NOV-1992; US-973341.
PR 29-JAN-1993; US-012990.
PR (ZONA-) ZONAGEN INC.
PA Harris JD, Hsu KT, Podolski JS;
PI WPI; 94-183156/22.
DR P-PSDB; R55197.
DR Use of zona pellucida proteins and antibodies - for inducing
PT reproducible transient infertility or permanent sterility in
PT female mammals
PS Claim 22; Page 81-83; I54pp; English.
CC A cDNA library was prepared from mRNA isolated from ovaries removed
CC from 5 week old rabbits. The lambda gt10 library was screened with
CC a porcine ZPC cDNA probe (Q65606). Two positive clones were sequenced
CC and were found to be identical except that one contained 4 extra
CC nucleotides at the 5'-end. The determined sequence was 75%
CC homologous to DNA coding for porcine ZPC.
SQ Sequence 1338 BP; 255 A; 445 C; 402 G; 236 T;

Query Match 2.0%; Score 43; DB 11; Length 1338;
Best Local Similarity 86.4%; Pred. No. 1.88e-07;
Matches 51; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 1260 aataaaatcagcttcacaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 1318
||||| ||| | ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 2114 AATAAATACTCAGAGCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2172

RESULT 10
ID Q47966 standard; cDNA; 1710 BP.
IC Q47966;
DT 25-MAR-1994 (first entry)
DE Rape acyl-ACP thioesterase (Clone pNL2) coding sequence.
KE Enzyme precursor; acyl-ACP thioesterase; plant; transformation;
KW fatty acid; yield; composition; seed oil; Brassica napus; ss.
FH Brassica napus.
FS Key Location/Qualifiers
FT CDS 169..1269
FT /*tag= a
FT /product= Enzyme precursor.
FT /note= "Has acyl-ACP thioesterase activity."
PN W09318158-A.
PD 16-SEP-1993.
PF 03-MAR-1993; G00432.
PR 03-MAR-1992; GB-004583.
PR (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.
PI Hellyer SA, Loader NM, Safford R, Slabas AR;
DR WPI; 93-303472/38.
DR P-PSDB; R41673.
PT cDNA encoding enzyme with acyl-ACP-thio:esterase activity -
PT useful for transforming plants, for altering properties e.g. seed
PT storage oil composition
PS Claim 1; Figure 1; 48pp; English.
CC The nucleotide sequence is useful for transforming plant cells and
CC altering that plants characteristics. Characteristics which may be
CC altered are (1) altering the level of acyl-ACP thioesterase
CC activity; (2) fatty acid yield and/or fatty acid composition and (3)
CC yield and/or composition of seed storage oil.
SQ Sequence 1710 BP; 534 A; 337 C; 364 G; 475 T;

```

CC The nucleotide sequence is useful for transforming plant cells and
CC altering that plants characteristics. Characteristics which may be
CC altered are (1) altering the level of acyl-ACP thioesterase
CC activity; (2) fatty acid yield and/or fatty acid composition and (3)
CC yield and/or composition of seed storage oil.
SQ Sequence 1710 BP; 534 A; 337 C; 364 G; 475 T;

RESULT 14
ID Q58663 standard; cDNA to mRNA; 699 BP.
AC Q58663;
DT 14-OCT-1994 (first entry)

RESULT	15	
AD	Q73729 standard; cDNA; 756 BP.	
IC	Q73729;	
DE	26-APR-1995 (first entry)	
DT	Rape abscission/dehiscence protein pSNC51 cDNA.	
DE	Rape; pSNC51; abscission; dehiscence; transgenic plant;	
KW	crop improvement; ss.	
KW	Brassica napus.	
OS	Location/Qualifiers	
FF	Key	
FT	CDS	15..458
FT	/*tag= a	
FT	polya_signal	633..638
FT	/*tag= b	
FT	/note= "putative polyA site"	
FN	W09423043-A.	
PD	13-OCT-1994.	
PP	31-MAR-1994; G00689.	
PPR	31-MAR-1993; GB-006726.	
PP	(NICK-) NICKERSON BIOCEM LTD.	
PFI	Coupe SA, Isaac PG, Roberts JA;	
DR	WPI: 94-333203/41.	
DR	P-PSDB; R60799.	
DR	Recombinant and isolated nucleic acids - encode enzymes and	
PT	proteins involved in plant abscission or dehiscence.	
PT	Disclosure; Fig 3; 39pp; English.	
CC	The cDNA encodes a rape (cv. Rafal) protein involved in abscission	
CC	or dehiscence, and can be used to manipulate plant abscission and	
CC	reduce/prevent pod dehiscence, especially to regulate abscission of	
CC	pollen from anthers, to generate male sterile plants and for hybrid	
CC	seed production.	
CC	Sequence	756 BP;
SC		240 A;
SC		175 C;
SC		118 G;
SC		223 T;

Search completed: Tue Aug 27 09:12:38 1996
Job time : 188 secs.

RESULT	15	
AD	Q73729 standard; cDNA; 756 BP.	
IC	Q73729;	
DE	26-APR-1995 (first entry)	
DT	Rape abscission/dehiscence protein pSNC51 cDNA.	
DE	Rape; pSNC51; abscission; dehiscence; transgenic plant;	
KW	crop improvement; ss.	
KW	Brassica napus.	
OS	Location/Qualifiers	
FF	Key	
FT	CDS	15..458
FT	/*tag= a	
FT	polya_signal	633..638
FT	/*tag= b	
FT	/note= "putative polyA site"	
FN	W09423043-A.	
PD	13-OCT-1994.	
PP	31-MAR-1994; G00689.	
PPR	31-MAR-1993; GB-006726.	
PP	(NICK-) NICKERSON BIOCEM LTD.	
PFI	Coupe SA, Isaac PG, Roberts JA;	
DR	WPI: 94-333203/41.	
DR	P-PSDB; R60799.	
DR	Recombinant and isolated nucleic acids - encode enzymes and	
PT	proteins involved in plant abscission or dehiscence.	
PT	Disclosure; Fig 3; 39pp; English.	
CC	The cDNA encodes a rape (cv. Rafal) protein involved in abscission	
CC	or dehiscence, and can be used to manipulate plant abscission and	
CC	reduce/prevent pod dehiscence, especially to regulate abscission of	
CC	pollen from anthers, to generate male sterile plants and for hybrid	
CC	seed production.	
CC	Sequence	756 BP;
SC		240 A;
SC		175 C;
SC		118 G;
SC		223 T;

WAVELENGTH (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.

Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.

Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 09:13:00 1996; MasPar time 847.32 Seconds
921.604 Million cell updates/sec

Tabular output not generated.

Title: >US-08-231-565A-26

Description: (1-2172) from US08231565A.seq

Perfect Score: 2172

N.A. Sequence: 1 GTGACGGCCATTACCAATC.....AAAAAAAAAAAAAAAAAAAA 2172

Comp: CACTCGCCGGTATGCTAG.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 514334 seqs, 179763086 bases x 2

Post-processing: Minimum Match 04

Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:STS2
93:STS3 94:STS4 95:STS5 96:STS6 97:STS7
EST-STS-TWO
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100
108:EST101 109:EST110 110:EST111 111:EST112 112:EST113
113:EST114 114:EST115 115:EST116 116:EST117 117:EST118
118:EST119 119:EST120 120:EST121 121:EST122 122:EST123
123:EST124 124:EST125 125:EST126 126:EST127
127:EST128 128:EST129 129:EST130 130:EST131 131:EST132
132:EST133 133:EST134 134:EST135 135:EST136 136:EST137
137:EST138 138:EST139 139:EST140 140:EST141 141:EST142
142:EST143 143:EST144 144:EST145 145:EST146 146:EST147
147:EST148 148:EST149 149:EST150 150:EST151 151:EST152
152:EST153 153:EST154 154:EST155 155:EST156 156:EST157
157:EST158 158:EST159 159:EST160 160:EST161 161:EST162
162:EST163 163:EST164 164:EST165 165:EST166 166:EST167
167:EST168 168:EST169 169:EST170 170:EST171 171:EST172
172:EST173 173:EST174 174:EST175 175:EST176 176:EST177
177:EST178 178:EST179 179:EST180 180:EST181 181:EST182
182:EST183 183:EST184 184:EST185 185:EST186 186:EST187
187:EST188 188:EST189 189:EST190 190:EST191 191:EST192
192:EST193 193:EST194 194:EST195 195:EST196 196:EST197
197:EST198 198:EST199 199:EST200 200:EST201 201:EST202
202:EST203 203:EST204 204:EST205 205:EST206 206:EST207
207:EST208 208:EST209 209:EST210 210:EST211 211:EST212
212:EST213 213:EST214 214:EST215 215:EST216 216:EST217
217:EST218 218:EST219 219:EST220 220:EST221 221:EST222
222:EST223 223:EST224 224:EST225 225:EST226 226:EST227
227:EST228 228:EST229 229:EST230 230:EST231 231:EST232
232:EST233 233:EST234 234:EST235 235:EST236 236:EST237
237:EST238 238:EST239 239:EST240 240:EST241 241:EST242
242:EST243 243:EST244 244:EST245 245:EST246 246:EST247
247:EST248 248:EST249 249:EST250 250:EST251 251:EST252
252:EST253 253:EST254 254:EST255 255:EST256 256:EST257
257:EST258 258:EST259 259:EST260 260:EST261 261:EST262
262:EST263 263:EST264 264:EST265 265:EST266 266:EST267
267:EST268 268:EST269 269:EST270 270:EST271 271:EST272
272:EST273 273:EST274 274:EST275 275:EST276 276:EST277
277:EST278 278:EST279 279:EST280 280:EST281 281:EST282
282:EST283 283:EST284 284:EST285 285:EST286 286:EST287
287:EST288 288:EST289 289:EST290 290:EST291 291:EST292
292:EST293 293:EST294 294:EST295 295:EST296 296:EST297
297:EST298 298:EST299 299:EST300 300:EST301 301:EST302
302:EST303 303:EST304 304:EST305 305:EST306 306:EST307
307:EST308 308:EST309 309:EST310 310:EST311 311:EST312
312:EST313 313:EST314 314:EST315 315:EST316 316:EST317
317:EST318 318:EST319 319:EST320 320:EST321 321:EST322
322:EST323 323:EST324 324:EST325 325:EST326 326:EST327
327:EST328 328:EST329 329:EST330 330:EST331 331:EST332
332:EST333 333:EST334 334:EST335 335:EST336 336:EST337
337:EST338 338:EST339 339:EST340 340:EST341 341:EST342
342:EST343 343:EST344 344:EST345 345:EST346 346:EST347
347:EST348 348:EST349 349:EST350 350:EST351 351:EST352
352:EST353 353:EST354 354:EST355 355:EST356 356:EST357
357:EST358 358:EST359 359:EST360 360:EST361 361:EST362
362:EST363 363:EST364 364:EST365 365:EST366 366:EST367
367:EST368 368:EST369 369:EST370 370:EST371 371:EST372
372:EST373 373:EST374 374:EST375 375:EST376 376:EST377
377:EST378 378:EST379 379:EST380 380:EST381 381:EST382
382:EST383 383:EST384 384:EST385 385:EST386 386:EST387
387:EST388 388:EST389 389:EST390 390:EST391 391:EST392
392:EST393 393:EST394 394:EST395 395:EST396 396:EST397
397:EST398 398:EST399 399:EST400 400:EST401 401:EST402
402:EST403 403:EST404 404:EST405 405:EST406 406:EST407
407:EST408 408:EST409 409:EST410 410:EST411 411:EST412
412:EST413 413:EST414 414:EST415 415:EST416 416:EST417
417:EST418 418:EST419 419:EST420 420:EST421 421:EST422
422:EST423 423:EST424 424:EST425 425:EST426 426:EST427
427:EST428 428:EST429 429:EST430 430:EST431 431:EST432
432:EST433 433:EST434 434:EST435 435:EST436 436:EST437
437:EST438 438:EST439 439:EST440 440:EST441 441:EST442
442:EST443 443:EST444 444:EST445 445:EST446 446:EST447
447:EST448 448:EST449 449:EST450 450:EST451 451:EST452
452:EST453 453:EST454 454:EST455 455:EST456 456:EST457
457:EST458 458:EST459 459:EST460 460:EST461 461:EST462
462:EST463 463:EST464 464:EST465 465:EST466 466:EST467
467:EST468 468:EST469 469:EST470 470:EST471 471:EST472
472:EST473 473:EST474 474:EST475 475:EST476 476:EST477
477:EST478 478:EST479 479:EST480 480:EST481 481:EST482
482:EST483 483:EST484 484:EST485 485:EST486 486:EST487
487:EST488 488:EST489 489:EST490 490:EST491 491:EST492
492:EST493 493:EST494 494:EST495 495:EST496 496:EST497
497:EST498 498:EST499 499:EST500 500:EST501 501:EST502
502:EST503 503:EST504 504:EST505 505:EST506 506:EST507
507:EST508 508:EST509 509:EST510 510:EST511 511:EST512
512:EST513 513:EST514 514:EST515 515:EST516 516:EST517
517:EST518 518:EST519 519:EST520 520:EST521 521:EST522
522:EST523 523:EST524 524:EST525 525:EST526 526:EST527
527:EST528 528:EST529 529:EST530 530:EST531 531:EST532
532:EST533 533:EST534 534:EST535 535:EST536 536:EST537
537:EST538 538:EST539 539:EST540 540:EST541 541:EST542
542:EST543 543:EST544 544:EST545 545:EST546 546:EST547
547:EST548 548:EST549 549:EST550 550:EST551 551:EST552
552:EST553 553:EST554 554:EST555 555:EST556 556:EST557
557:EST558 558:EST559 559:EST560 560:EST561 561:EST562
562:EST563 563:EST564 564:EST565 565:EST566 566:EST567
567:EST568 568:EST569 569:EST570 570:EST571 571:EST572
572:EST573 573:EST574 574:EST575 575:EST576 576:EST577
577:EST578 578:EST579 579:EST580 580:EST581 581:EST582
582:EST583 583:EST584 584:EST585 585:EST586 586:EST587
587:EST588 588:EST589 589:EST590 590:EST591 591:EST592
592:EST593 593:EST594 594:EST595 595:EST596 596:EST597
597:EST598 598:EST599 599:EST600 600:EST601 601:EST602
602:EST603 603:EST604 604:EST605 605:EST606 606:EST607
607:EST608 608:EST609 609:EST610 610:EST611 611:EST612
612:EST613 613:EST614 614:EST615 615:EST616 616:EST617
617:EST618 618:EST619 619:EST620 620:EST621 621:EST622
622:EST623 623:EST624 624:EST625 625:EST626 626:EST627
627:EST628 628:EST629 629:EST630 630:EST631 631:EST632
632:EST633 633:EST634 634:EST635 635:EST636 636:EST637
637:EST638 638:EST639 639:EST640 640:EST641 641:EST642
642:EST643 643:EST644 644:EST645 645:EST646 646:EST647
647:EST648 648:EST649 649:EST650 650:EST651 651:EST652
652:EST653 653:EST654 654:EST655 655:EST656 656:EST657
657:EST658 658:EST659 659:EST660 660:EST661 661:EST662
662:EST663 663:EST664 664:EST665 665:EST666 666:EST667
667:EST668 668:EST669 669:EST670 670:EST671 671:EST672
672:EST673 673:EST674 674:EST675 675:EST676 676:EST677
677:EST678 678:EST679 679:EST680 680:EST681 681:EST682
682:EST683 683:EST684 684:EST685 685:EST686 686:EST687
687:EST688 688:EST689 689:EST690 690:EST691 691:EST692
692:EST693 693:EST694 694:EST695 695:EST696 696:EST697
697:EST698 698:EST699 699:EST700 700:EST701 701:EST702
702:EST703 703:EST704 704:EST705 705:EST706 706:EST707
707:EST708 708:EST709 709:EST710 710:EST711 711:EST712
712:EST713 713:EST714 714:EST715 715:EST716 716:EST717
717:EST718 718:EST719 719:EST720 720:EST721 721:EST722
722:EST723 723:EST724 724:EST725 725:EST726 726:EST727
727:EST728 728:EST729 729:EST730 730:EST731 731:EST732
732:EST733 733:EST734 734:EST735 735:EST736 736:EST737
737:EST738 738:EST739 739:EST740 740:EST741 741:EST742
742:EST743 743:EST744 744:EST745 745:EST746 746:EST747
747:EST748 748:EST749 749:EST750 750:EST751 751:EST752
752:EST753 753:EST754 754:EST755 755:EST756 756:EST757
757:EST758 758:EST759 759:EST760 760:EST761 761:EST762
762:EST763 763:EST764 764:EST765 765:EST766 766:EST767
767:EST768 768:EST769 769:EST770 770:EST771 771:EST772
772:EST773 773:EST774 774:EST775 775:EST776 776:EST777
777:EST778 778:EST779 779:EST780 780:EST781 781:EST782
782:EST783 783:EST784 784:EST785 785:EST786 786:EST787
787:EST788 788:EST789 789:EST790 790:EST791 791:EST792
792:EST793 793:EST794 794:EST795 795:EST796 796:EST797
797:EST798 798:EST799 799:EST800 800:EST801 801:EST802
802:EST803 803:EST804 804:EST805 805:EST806 806:EST807
807:EST808 808:EST809 809:EST810 810:EST811 811:EST812
812:EST813 813:EST814 814:EST815 815:EST816 816:EST817
817:EST818 818:EST819 819:EST820 820:EST821 821:EST822
822:EST823 823:EST824 824:EST825 825:EST826 826:EST827
827:EST828 828:EST829 829:EST830 830:EST831 831:EST832
832:EST833 833:EST834 834:EST835 835:EST836 836:EST837
837:EST838 838:EST839 839:EST840 840:EST841 841:EST842
842:EST843 843:EST844 844:EST845 845:EST846 846:EST847
847:EST848 848:EST849 849:EST850 850:EST851 851:EST852
852:EST853 853:EST854 854:EST855 855:EST856 856:EST857
857:EST858 858:EST859 859:EST860 860:EST861 861:EST862
862:EST863 863:EST864 864:EST865 865:EST866 866:EST867
867:EST868 868:EST869 869:EST870 870:EST871 871:EST872
872:EST873 873:EST874 874:EST875 875:EST876 876:EST877
877:EST878 878:EST879 879:EST880 880:EST881 881:EST882
882:EST883 883:EST884 884:EST885 885:EST886 886:EST887
887:EST888 888:EST889 889:EST890 890:EST891 891:EST892
892:EST893 893:EST894 894:EST895 895:EST896 896:EST897
897:EST898 898:EST899 899:EST900 900:EST901 901:EST902
902:EST903 903:EST904 904:EST905 905:EST906 906:EST907
907:EST908 908:EST909 909:EST910 910:EST911 911:EST912
912:EST913 913:EST914 914:EST915 915:EST916 916:EST917
917:EST918 918:EST919 919:EST920 920:EST921 921:EST922
922:EST923 923:EST924 924:EST925 925:EST926 926:EST927
927:EST928 928:EST929 929:EST930 930:EST931 931:EST932
932:EST933 933:EST934 934:EST935 935:EST936 936:EST937
937:EST938 938:EST939 939:EST940 940:EST941 941:EST942
942:EST943 943:EST944 944:EST945 945:EST946 946:EST947
947:EST948 948:EST949 949:EST950 950:EST951 951:EST952
952:EST953 953:EST954 954:EST955 955:EST956 956:EST957
957:EST958 958:EST959 959:EST960 960:EST961 961:EST962
962:EST963 963:EST964 964:EST965 965:EST966 966:EST967
967:EST968 968:EST969 969:EST970 970:EST971 971:EST972
972:EST973 973:EST974 974:EST975 975:EST976 976:EST977
977:EST978 978:EST979 979:EST980 980:EST981 981:EST982
982:EST983 983:EST984 984:EST985 985:EST986 986:EST987
987:EST988 988:EST989 989:EST990 990:EST991 991:EST992
992:EST993 993:EST994 994:EST995 995:EST996 996:EST997
997:EST998 998:EST999 999:EST1000 1000:EST1001 1001:EST1002
1002:EST1003 1003:EST1004 1004:EST1005 1005:EST1006 1006:EST1007
1007:EST1008 1008:EST1009 1009:EST1010 1010:EST1011 1011:EST1012
1012:EST1013 1013:EST1014 1014:EST1015 1015:EST1016 1016:EST1017
1017:EST1018 1018:EST1019 1019:EST1020 1020:EST1021 1021:EST1022
1022:EST1023 1023:EST1024 1024:EST1025 1025:EST1026 1026:EST1027
1027:EST1028 1028:EST1029 1029:EST1030 1030:EST1031 1031:EST1032
1032:EST1033 1033:EST1034 1034:EST1035 1035:EST1036 1036:EST1037
1037:EST1038 1038:EST1039 1039:EST1040 1040:EST1041 1041:EST1042
1042:EST1043 1043:EST1044 1044:EST1045 1045:EST1046 1046:EST1047
1047:EST1048 1048:EST1049 1049:EST1050 1050:EST1051 1051:EST1052
1052:EST1053 1053:EST1054 1054:EST1055 1055:EST1056 1056:EST1057
1057:EST1058 1058:EST1059 1059:EST1060 1060:EST1061 1061:EST1062
1062:EST1063 1063:EST1064 1064:EST1065 1065:EST1066 1066:EST1067
1067:EST1068 1068:EST1069 1069:EST1070 1070:EST1071 1071:EST1072
1072:EST1073 1073:EST1074 1074:EST1075 1075:EST1076 1076:EST1077
1077:EST1078 1078:EST1079 1079:EST1080 1080:EST1081 1081:EST1082
1082:EST1083 1083:EST1084 1084:EST1085 1085:EST1086 1086:EST1087
1087:EST1088 1088:EST1089 1089:EST1090 1090:EST1091 1091:EST1092
1092:EST1093 1093:EST1094 1094:EST1095 1095:EST1096 1096:EST1097
1097:EST1098 1098:EST1099 1099:EST1100 1100:EST1101 1101:EST1102
1102:EST1103 1103:EST1104 1104:EST1105 1105:EST1106 1106:EST1107
1107:EST1108 1108:EST1109 1109:EST1110 1110:EST1111 1111:EST1112
1112:EST1113 1113:EST1114 1114:EST1115 1115:EST1116 1116:EST1117
1117:EST1118 1118:EST1119 1119:EST1120 1120:EST1121 1121:EST1122
1122:EST1123 1123:EST1124 1124:EST1125 1125:EST1126 1126:EST1127
1127:EST1128 1128:EST1129 1129:EST1130 1130:EST1131 1131:EST1132
1132:EST1133 1133:EST1134 1134:EST1135 1135:EST1136 1136:EST1137
1137:EST1138 1138:EST1139 1139:EST1140 1140:EST1141 1141:EST1142
1142:EST1143 1143:EST1144 1144:EST1145 1145:EST1146 1146:EST1147
1147:EST1148 1148:EST1149 1149:EST1150 1150:EST1151 1151:EST1152
1152:EST1153 1153:EST1154 1154:EST1155 1155:EST1156 1156:EST1157
1157:EST1158 1158:EST1159 1159:EST1160 1160:EST1161 1161:EST1162
1162:EST1163 1163:EST1164 1164:EST1165 1165:EST1166 1166:EST1167
1167:EST1168 1168:EST1169 1169:EST1170 1170:EST1171 1171:EST1172
1172:EST1173 1173:EST1174 1174:EST1175 1175:EST1176 1176:EST1177
1177:EST1178 1178:EST1179 1179:EST1180 1180:EST1181 1181:EST1182
1182:EST1183 1183:EST1184 1184:EST1185 1185:EST1186 1186:EST1187
1187:EST1188 1188:EST1189 1189:EST1190 1190:EST1191 1191:EST1192
1192:EST1193 1193:EST1194 1194:EST1195 1195:EST1196 1196:EST1197
1197:EST1198 1198:EST1199 1199:EST1200 1200:EST1201 1201:EST1202
1202:EST1203 1203:EST1204 1204:EST1205 1205:EST1206 1206:EST1207
1207:EST1208 1208:EST1209 1209:EST1210 1210:EST1211 12

LOCUS N28728 621 bp mRNA EST 04-JAN-1996
DEFINITION yk67a12.r1 Homo sapiens cDNA clone 266782 5' similar to
SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte
lineage - human ;
ACCESSION N28728
NID g1146964
KEYWORDS EST.
SOURCE human clone-266782 primer=T7 library=Soares melanocyte 2Nblm
vector=pT73D (Pharmacia) with a modified polylinker host=PH10B
(ampicillin resistant) Reitel=Not I Reite2=Eco RI Male. 1st strand
cDNA was primed with a Not I - oligo (dT) primer
[5'-TGTACCAATCTGAAGTGGAGCGCGCGAGTGTGTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 621)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

High quality sequence stops: 398

Source: IMAGE Consortium, LIND

This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.lind.gov) for further information.

NCBI gi: 1146964

FEATURES Location/Qualifiers

source 1..621

/organism="Homo sapiens"

/clone="266782"

/note="human"

<1..>621

BASE COUNT 125 a 170 c 173 g 145 t 8 others

ORIGIN

Query Match 25.0%; Score 542; DB 114; Length 621;

Best Local Similarity 96.2%; Pred. No. 0.00e+00;

Matches 604; Conservative 0; Mismatches 15; Indels 9; Gaps 9;

Db 1 ctgagagctacctccctgagctgaaggtccagatgccagctcaatcatgtctacgg 60

|||||

Qy 1321 CTGAGAGCTACCTATCCCTGAGCGCTGAAGCTCCAGATCCGAGCTCAATCATGTCTACGG 1380

|||||

Db 61 aaagtattacagttccctggccccctgctggatgtagcagccaccttaagctgtga 120

|||||

Qy 1381 AAGATATTACAGTTCCCTGGGCCCTCTGCTGATGCTAGCAGCCACTTAAGCTGTGCTCA 1440
Db 121 aagagcaagtccccctgattgtgtctgtatcatgatatgggtctcttttccgtcaccctgg 180
|||||
Qy 1441 AGAGACAAGTCCCCCTGGATTGTGTCTGATCGATATGGTCTCTTTTCCGCTCACCTGG 1500
Db 181 acattgtccagatnaintgaaagtccagatcctgcaggtgtgctgncsggtgagggg 240
|||||
Qy 1501 ACATGTGCCAGGTAT-TGAAGTGGCCAGATCTCCAGGCTGTCCGCTCCGCTGAGGGG 1559
Db 241 gatcatttgagctgactgtctgccaaggcggtgcccagaagaagcctcatgtgag 300
|||||
Qy 1560 GATGCAATTCAGCTGCTGCTGCTGCCAAGGGGGCTGCCAAGGAAGCCTGATGGAG 1619
|||||
Db 301 atctcatcgccaggtgccagccctgccagcggtgtgc-agcctgtgtaccagc 359
|||||
Qy 1620 ATCTCATCGCCAGGGTGGCAGCCCTGCCCCAGCGGTGCGCAGCGCTGTGCTACCCAGC 1679
Db 360 ccagctgccagctggttctgcaccagatactgaaggtggctcggggacatactgcctc 419
|||||
Qy 1680 CCAGCTGCCAGCTGCTTCTGCACCATATCTCAAGGGTGGCTCCGGGACATATCTGCCTC 1739
Db 420 aatgtctctggtgatataccagcagctggcagtggtcagcagcagcagcttatcatgct 479
|||||
Qy 1740 AATGTGCTCTGGCTGATACCAAGCGCTGGCAGTGGTTCAGCACCAGCTTATCATGCT 1799
Db 480 ggtcaagaagcaggttggnaaggttcctgtatccgtgggcancttgggtgtgtgat 539
|||||
Qy 1800 GGTCAAGACAGCGCTTGGCGAGGTTCCTGCTATC-CTGGGCACTTCTGCTGTGTGAT 1858
Db 540 ggctgtggcctttcatctctgntataggg-ca-acttatgaagca-actt-tnogt 595
|||||
Qy 1859 GGCTGTGCTCTTGCATCTCTGATATATAGCGCAGACTTATGAAGCAAGACTTCTCGGT 1918
Db 596 accc-agttgc-acatagcagcagcagc 621
|||||
Qy 1919 ACCCGAGTGGCACATAGCAGCAGTCAAC 1946

RESULT 2

ID HS728262 standard; RNA; EST; 621 BP.

AC N28728;

DT 05-JAN-1996 (Rel. 46, Created)

DT 05-JAN-1996 (Rel. 46, Last updated, Version 1)

DE yk67a12.r1 Homo sapiens cDNA clone 266782 5' similar to

DE SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte

DE lineage - human ;

KW EST.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.

RN [1]

RP 1-621

RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,

RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,

RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;

RT "The WashU-Merck EST Project";

RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University

CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,

CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@watson.wustl.edu High quality sequence stops: 398 Source: IMAGE

CC Consortium, LIND This clone is available royalty-free through LIND

Aug 27 09:14

US-08-231-565A-26.rst

5

CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
information. NCBI gi: 1146964
FH Key Location/Qualifiers

FT source 1..621
FT /organism="Homo sapiens"
FT /clone="266782"
FT /note="human"
FT mRNA <1..>621
SQ Sequence 621 BP; 125 A; 170 C; 173 G; 145 T; 8 other;

Query Match 25.08; Score 542; DB 136; Length 621;
Best Local Similarity 96.2%; Pred. No. 0.00e+00;
Matches 604; Conservative 0; Mismatches 15; Indels 9; Gaps 9;

Db 1 cttagagagctacctccctgagcctgaaggtccagatgccagctcaatcatgtctacgg 60
|||||
Qy 1321 CTAGAGAGCTACTATCCCTGAGCTGAGGTCCAGATGCCAGCTCATCATCTGTACGG 1380
|||||
Db 61 aaagtattacaggttccctggcccccctgctggatggtacagccacttaaggtcgtga 120
|||||
Qy 1381 AAAGTATTACAGGTTCCCTGGGCCCCCTGCTGGATGGTACAGCCACTTAAGGCTGGTGA 1440
|||||
Db 121 agagacaagtcceccctggatgtgtctgatcatgatgttcttcttccctcaccctgg 180
|||||
Qy 1441 AGAGACAAGTCCCTGGATGTTGTTCTGTATCATGATATGTTCTTTCCTCCGTCACCTGG 1500
|||||
Db 181 acattgtccaggtatnattgaaggtccagatccctgcaggtgtgcgtncctggagggg 240
|||||
Qy 1501 ACATTGTCCAGGGTAT-TGAAGTCCGAGATCTTCAGGCTGTGGCTCGCGTGGAGGG 1559
|||||
Db 241 gatcattgagctgaactgtgctgccaagcggggtgcccaaggaagcctgcattgag 300
|||||
Qy 1560 GATGCAATTGACGTGCTGTGCTGCCAGGGGGGCTGCCAAGGAGCCCTGCATGAG 1619
|||||
Db 301 atctcatgccaggtgccagcccccctgccagcggctgtgc-agcctgtgtaccagc 359
|||||
Qy 1620 ATCTATGCCAGGTGCCAGGCCCTGCCAGCGGCTGTGCCAGCCTGTCTACCCAGC 1679
|||||
Db 360 ccagcctgccagctggtctgcaccagatctgaaggtggtctcggggacactactgcctc 419
|||||
Qy 1680 CCAGCCTGCCAGCTGTTCTTGACCAAGATCTGAAGGGTGGCTCGGGGACATATCGCTC 1739
|||||
Db 420 aatgtgtctgtgtatataccaaagcctggcagtggtcagcaaccagcttatcatgct 479
|||||
Qy 1740 AATGTGTCTGTGGCTGATACCAAGCCTGGCAGTGGTGTACGACCCAGCTTATCATGCT 1799
|||||
Db 480 ggtcaagaagcaggccttgnaaggttccgtgatccgtgggcancttgcgtggtgtgat 539
|||||
Qy 1800 GGTCAAGACAGAGCCCTGGCAGGTGCCGCTGATC-GTGGGCACTTGTGCTGCTGTGAT 1858
|||||
Db 540 ggctgtggcctttcatctcgtatataagc-ca-acttatgaagcaa-actt-tncgt 595
|||||
Qy 1859 GGCTGTGGCTTGTGCTCTGTATATAGCGCAGACTTATGAAGCAAGACTTCTCGCT 1918
|||||
Db 596 accc-agttgc-acatagcagcagaaac 621
|||||
Qy 1919 ACCCCAGTTGCCACATAGCAGCAGTCA 1946
|||||

RESULT 3
LOCUS N28736 638 bp mRNA EST 04-JAN-1996
DEFINITION yx67cl2.r1 Homo sapiens cDNA clone 266806 5' similar to
SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte

Aug 27 09:14

US-08-231-565A-26.rst

6

Lineage - human ;
ACCESSION N28736
NID g1146972
KEYWORDS EST.
SOURCE human

human clone=266806 primer=T7 library=Soares melanocyte 2N8IM
vector=pfT73D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) Rsite=Not I Rsite2=Eco RI Male. 1st strand
cDNA was primed with a Not I - oligo(dT) primer
[5'-TCCTACCAATCTGAAGTGGGCGCGGCGAGTCTTTTTTTTTTTTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). Library
constructed by Berto Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 638)

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 368
Source: IMAGE Consortium, LINL
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1146972

Location/Qualifiers

FEATURES

source

1..638
/organism="Homo sapiens"
/clone="266806"
/note="human"

mRNA

BASE COUNT 133 a 169 c 173 g 153 t 10 others

ORIGIN

Query Match 22.8%; Score 496; DB 114; Length 638;
Best Local Similarity 95.5%; Pred. No. 0.00e+00;
Matches 547; Conservative 0; Mismatches 20; Indels 6; Gaps 6;

Db 1 cttagagagctacctccctgagcctgaaggtccagatgccagctcaatcatgtctacgg 60
|||||

Qy 1321 CTAGAGAGCTACTATCCCTGAGCTGAGGTCCAGATGCCAGCTCAATCATCTGTACGG 1380
|||||

Db 61 aaagtattacaggttccctggcccccctgctggatggtacagccacttaaggtcgtga 120
|||||

Qy 1381 AAAGTATTACAGGTTCCCTGGGCCCCCTGCTGGATGGTGTACAGCCACTTAAGGCTGGTGA 1440
|||||

Db 121 agagacaagtcceccctggatgtgtctgatcatgatgttcttcttccctcaccctgg 180
|||||

```

QY 1441 ACAGCAAGTCCCTGGATTGTTCTGTATCGATATGGTTCTTTTCCTGCACCCCTGG 1500
|||||
Db 181 acattgtccaggttatttgaagtcgagactctgcaggtgtgctgctnccgtgagggg 240
|||||
QY 1501 ACATTTGTCAGGGTATT-GAAAGTCGCGATCTCTGCAGGCTGTGCGCTGCGGTGAGGGG 1559
|||||
Db 241 qatgcattgaactgaactgtgctgccaagcgagctgccaaggaagcctgcatgag 300
|||||
QY 1560 GATGCATTTGAGCTGACTGTGTCTGCTGCAAGGGGGCTGCCAAGAAAGCCTGCATGCGAG 1619
|||||
Db 301 atctcatcgccaggtgcccagccccctgcccagcggtgtgc-agcctgtgctaccaccg 359
|||||
QY 1620 ATCTCATGCGCAGGTCGCCAGGCCCTGCCAGCGGCTGTGCCAGGCTGTCTTACCAGC 1679
|||||
Db 360 ccagctgcagctgggttgcaccagataactgaaggtggtgctgctgggaacaactgct 419
|||||
QY 1680 CCAGCTGCCAGCTGGTTCTGCACAGACTGAGAGGCTGGCTCGGGCA-CATACCTGCT 1738
|||||
Db 420 caatgtctctggtgataacaacagcctggcagtggtgcagcaccagcttatcatgc 479
|||||
QY 1739 CAATGTCTCTGGCTGATACCAA-CAGCCTGGCAGTGGTCAGCACCAGCTTATCATGC 1797
|||||
Db 480 ctggtcaagaacaggtctgnaaggtccgctgactgctggggactctgctggtgta 539
|||||
QY 1798 CTGGTCAAGACGGCTTGGCGAGGTTCCGCTGATCGTGGG-CATCTGCTGCTGTTG 1856
|||||
Db 540 atgctgtgg-ccttgnacccttataaaatagg 571
|||||
QY 1857 ATGGCTGTGGCTCTGCATCTCTGATATAGG 1889
|||||

```

RESULT 4

```

ID HS736264 standard; RNA; EST; 638 BP.
AC N28736;
DT 05-JAN-1996 (Rel. 46, Created)
DE yx67c12.rl Homo sapiens cDNA clone 266806 5' similar to
DE SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte
DE lineage - human ;.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-638
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@wustl.wustl.edu High quality sequence stops: 368 Source: IMAGE
CC Consortium, LNL This clone is available royalty-free through LNL
CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further
CC information. NCBI gi: 1146972
FH Key Location/Qualifiers
FT source
FT 1..638
FT /organism="Homo sapiens"
FT /clone="266806"

```

```

FT mRNA /note="human"
SQ Sequence 638 BP; 133 A; 169 C; 173 G; 153 T; 10 other;

Query Match 22.8%; Score 496; DB 136; Length 638;
Best Local Similarity 95.5%; Pred. No. 0.00e+00;
Matches 547; Conservative 0; Mismatches 20; Indels 6; Gaps 6;

Db 1 ctgagagctactatccctgagcctgaaggtccagatgccagctcaatcatgtctacgg 60
|||||
QY 1321 CTAGAGAGCTACCTATCCCTGAGCCTGAAGTCCAGATGCCAGCTCAATCATGTCTACGG 1380
|||||
Db 61 aaagtattacaggttccctggccccctgctggatggtacagccaccttaaggtggtga 120
|||||
QY 1381 AAGATTATTACAGGTTCCCTGGGCCCCCTGCTGGATGTGTACAGCCACCTTAAGGCTGTCA 1440
|||||
Db 121 agagcaagtcctccctggattgtgtctgcatgatatggttctcttccgtcacctgg 180
|||||
QY 1441 AGAGCAAGTCCCTCGATTGTGTCTGTATCGATATGGTTCTTTTCCGTACCCCTGG 1500
|||||
Db 181 acattgtccaggttatttgaagtcgagatccctgcaggtgtgctnccgtgagggg 240
|||||
QY 1501 ACATTTGTCAGGGTATT-GAAAGTCCGAGATCCTGCAGGCTGCGCGTGAGGGG 1559
|||||
Db 241 gatgcattgagctgaactgtgctctgccaagcggtgctgccaaggaagcctgcatgag 300
|||||
QY 1560 GATGCATTTGAGCTGACTGTGTCTGCTGCCAAGCGGGCTGCCAAGGAAGCCTGCATGGAG 1619
|||||
Db 301 atctcatcgccaggtgcccagccccctgcccagcggtgtgc-agcctgtgctaccaccg 359
|||||
QY 1620 ATCTCATGCGCAGGTCGCCAGGCCCTGCCAGCGGCTGTGCCAGGCTGTGCTTACCAGC 1679
|||||
Db 360 ccagctgcagctgggttgcaccagataactgaaggtggtgctgagggaacatactgct 419
|||||
QY 1680 CCAGCTGCCAGCTGGTTCTGCACAGACTGAGAGGCTGGCTCGGGCA-CATACCTGCT 1738
|||||
Db 420 caatgtctctggtgataacaacagcctggcagtggtgcagcaccagcttatcatgc 479
|||||
QY 1739 CAATGTCTCTGGCTGATACCAA-CAGCCTGGCAGTGGTCAGCACCAGCTTATCATGC 1797
|||||
Db 480 ctggtcaagaacaggtctgnaaggttccgctgactgctggggactctgctggtgta 539
|||||
QY 1798 CTGGTCAAGACGGCTTGGCGAGGTTCCGCTGATCGTGGG-CATCTGCTGCTGTTG 1856
|||||
Db 540 atgctgtgg-ccttgnacccttataaaatagg 571
|||||
QY 1857 ATGGCTGTGGCTCTGCATCTCTGATATAGG 1889
|||||

```

RESULT 5

```

ID HS746262 standard; RNA; EST; 546 BP.
AC N28746;
DT 05-JAN-1996 (Rel. 46, Created)
DE yx67f11.rl Homo sapiens cDNA clone 266829 5' similar to
DE SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte
DE lineage - human ;.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-546
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

```

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.

CC Contact: Willson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@watson.wustl.edu High quality sequence stops: 377 Source: IMAGE
CC Consortium, LNL This clone is available royalty-free through LNL
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1146982

FH Key Location/Qualifiers

FT source

1..546

/organism="Homo sapiens"

/clone="266829"

/note="human"

FT mRNA

<1..>546

FT Sequence 546 bp; 114 A; 145 C; 158 G; 120 T; 9 other;

Query Match

19.6%; Score 426; DB 136; Length 546;

Best Local Similarity 94.8%; Pred. No. 0.00e+00;

Matches 507; Conservative 0; Mismatches 16; Indels 12; Gaps 11;

Db 2 catntacagagtgggtggagacacagctagagagctacattccctgagcctgaagtc 61

|| |||||

Qy 1294 CAACTACAGAGTGGGTGGAGACACAGCTAGAGCTACCTATCCCTGAGCTGAAGTTC 1353

Db 62 cagatgccagctcaatcatgtctacggaaagtattacaggttccctgggcccctgtcgg 121

|||||

Qy 1354 CAGATGCCAGCTCAATCATCTCTACGAAAGTATTACAGGTTCCCTGGGGCCCTGCTGG 1413

Db 122 atgtcacagccacctaaagctggtgaagagacaagtcaccttgattgtctgtatc 181

|||||

Qy 1414 ATGTACAGCCACCTTAAGCTGTGTGAAGACAGCAAGTCCCTGTGATTTGTGTGTATC 1473

Db 182 gatatgttcttttcgtcacctggacattgtccaggttnatntgaagtgcagatc 241

|||||

Qy 1474 GATATGTTTCCTTTCCGTACCCCTGGACATTTGTCCAGGTAT-TGAAGAATGCCAGATC 1532

Db 242 ctgaggctgtcgtcccggtgaggggatgcatcttgagctgactgtcctccaagtc 301

|||||

Qy 1533 CTGCAGGCTGTCCCTCCGTGAGGGGATGCAATTTGAGCTGACTGTCTCTGCCAAGGC 1592

Db 302 gggctgcccaagaagctgcattgagatctcatgccaggtgccagccctgccag 361

|||||

Qy 1593 GGGCTGCCAAGAGAGCTGCATGAGATCTCATGCCAGGGTGCAGCCCTGCCAG 1652

Db 362 cggctgtgc-acctgtcgtaccagccagctgacagctggttctgcaccagatc 420

|||||

Qy 1653 CGGCTGTGCCAGCTGTCTTACCAGCCAGCTGCAGCTGTCTGCACGAG-ATACT 1711

Db 421 gaagggtggctcggggaaacataactggctcaatgtctctctggtggatncaacagc 480

|||||

Qy 1712 GAAAGG-TGGCTCGGGG--ACATA-CTGCCTCAATGTCTCTGTGGCTG-ATACCAACAGC 1766

Db 481 ctggcagttgtnaagcaacccactttatcatgctggtgggnaagaacaggnctt 535

|||||

Qy 1767 CTGGCACT-GCTCA-GCA-CCCAGCTTATCATGCCCTGG-TCAAGAGAGCGCCTT 1817

RESULT 6

LOCUS N28746 546 bp mRNA EST 04-JAN-1996

DEFINITION yx67fll.1 Homo sapiens cDNA clone 266829 5' similar to
SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte
lineage - human ;

ACCESSION N28746

NID g1146982

KEYWORDS EST.

SOURCE

human clone=266829 primer=T7 library=Soares melanocyte 2NbhM
vector=pT73D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) RseI=Not I RseI2=Eco RI Male. 1st strand

cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGAGTTTTTTTTTTTTTTT-3'],

double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco

RI sites of a modified pT73 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal

foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.

Albino.

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;

Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;

Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;

Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 546)

AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,

Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,

Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,

Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and

Wilson,R.

THE WashU-Merck EST Project

Unpublished (1995)

CONTACT: Willson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 377

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1146982

FEATURES

source

1..546

/organism="Homo sapiens"

/clone="266829"

/note="human"

<1..>546

BASE COUNT

114 a 145 c 158 g 120 t 9 others

ORIGIN

Query Match

19.6%; Score 426; DB 114; Length 546;

Best Local Similarity 94.8%; Pred. No. 0.00e+00;

Matches 507; Conservative 0; Mismatches 16; Indels 12; Gaps 11;

Db 2 catntacagagtgggtggagacacagctagagagctacattccctgagcctgaagtc 61

|| |||||

Qy 1294 CAACTACAGAGTGGGTGGAGACACAGCTAGAGCTACCTATCCCTGAGCTGAAGTTC 1353

Db 62 cagatgccagctcaatcatgtctacggaaagtattacaggttccctgggcccctgtcgg 121

|||||

Qy 1354 CAGATGCCAGCTCAATCATCTCTACGAAAGTATTACAGGTTCCCTGGGGCCCTGCTGG 1413

Db 122 atggtacagccacccatgaagctggtgaagagacagcccccctggattgtgtctgtatc 181
|||||
Qy 1414 ATGGTACAGCCACCTTAAAGCTGGTGAAGAGACAAGTCCCCCTGGATTGTGTCTGTATC 1473
|||||
Db 182 gatatgttcccttttcgcgtcacccctggaacatgtccaggtnatntgaaagtgcagagac 241
|||||
Qy 1474 GATATGTTCTCTTTTCGCTCACCTGGACATTGTCCAGGGTAT-TCAAAGTGGCGAGATC 1532
|||||
Db 242 ctgcaggtgtgcgtcccgtgtagggggtgcatcttgagctgaactgtctcctccaaggg 301
|||||
Qy 1533 CTGCAGGCTGTCCCTCGCTGCGGGGATGCAATTGAGCTGACTGTCTCTCCCAAGGC 1592
|||||
Db 302 gggctgcccagaagccgtgcagatgcagatcctcatcgccaggggtgcagccccctgccag 361
|||||
Qy 1593 GGGCTGCCAAGGAGCTGTGATGAGATCTCATGCCAGGGTGGCAGCCCTGCCCG 1652
|||||
Db 362 cggctgtgc-aggctgtctaccagccagcgcctcagctggttctgcaccaggatct 420
|||||
Qy 1653 CGGCTGTGCCAGCTGTGTACCCAGCCAGCCTGCCAGCTGGTTCGACACGAG-ATACT 1711
|||||
Db 421 gaaaggggtcgtcggggaacataactggctcaatgtgtctcgtgctgatancacagc 480
|||||
Qy 1712 GAAAGGG-TGGCTCGGGG--ACATA-CTGCCCTCAATGTGTCTCTGGCTG-ATACCAACAGC 1766
|||||
Db 481 ctggcaggtgtnaagcaacccactttatcatgctggtggtggtggtggtggtggtggt 535
|||||
Qy 1767 CTGGCAGT-GGTCA-GCA-CCGAGCTTATCATGCTCTGG-TCAAGAAGCAGCGCTT 1817
|||||

RESULT 7

ID HS118284 standard; RNA; EST; 430 BP.
AC N23118;
DT 29-DEC-1995 (Rel. 46, Created)
DT 29-DEC-1995 (Rel. 46, Last updated, Version 1)
DE yx67cl2.s1 Homo sapiens cDNA clone 266806 3' similar to SP:A41234
DE A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-430
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RA "The WashU-Merck EST Project";
RT Unpublished.
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@wustl.wustl.edu High quality sequence stops: 288 Source: IMAGE
CC Consortium, LINL This clone is available royalty-free through LINL
CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further
CC information. NCBI gi: 1137268
FH Key Location/Qualifiers
FH source
FT 1..430
FT /organism="Homo sapiens"
FT /clone="266806"
FT /notes="human"
FT mRNA
FT <1..>430
SQ Sequence 430 BP; 114 A; 107 C; 115 G; 93 T; 1 other;

Query Match 19.3%; Score 420; DB 130; Length 430;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 427; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Db 1 caggctcgtgattatttatttcagtttaataagtagtctcccagggaagactgggggaatat 60
|||||
Cp 2132 CAGGCTCTGAGTATTTATTTTCAGTTAAATAGTAGTCTCCAGGGAAGACTGGGGAAAT 2073
|||||
Db 61 aggtgtttctgtcaactccaggagaaatcacagcatcatatgagagtactcagacctgtg 120
|||||
Cp 2072 AGGTGTTTCTGTAACCTCCAGGAAATCACAGCATCATATGAGAGTACTCAGACCTGCTG 2013
|||||
Db 121 cccactgaggagggggtattctcccaatgggacaagagcagaagatcggggtagacg 180
|||||
Cp 2012 CCCACTGAGGAGGGGGCTGTCTCACCATGGGACAAGAGCAGAGATGGGGGTAGACG 1953
|||||
Db 181 cagcagtgactgtctgtatgtgcaactgggtgacggagaagtctgtctcataagctc 240
|||||
Cp 1952 CAGCAGCTGACTGTCTGTATGTGGCAACTGGGTACGAGAGACTCTGCTTCTATAGTCT 1893
|||||
Db 241 ggcctatatatcagatgcagagcaagcaacacacacacacacacacacacacacacac 300
|||||
Cp 1892 GGCCTATATATAGAGATGCAAGGACACACGCCATCAACACCAAGATGCCACGAT 1833
|||||
Db 301 cagcgaacctgcccaaggctgtctctgaccagcatgataagctgggtgcttgacca 360
|||||
Cp 1832 CAGCGGAACCTGCCAAGGCTGCTTCTTGACAGGCATGATAAGCTGGTGTCT-GACCA 1774
|||||
Db 361 ctgcagggtnttggtatcagccagagacacattgagcagtagtgcgccagccacct 420
|||||
Cp 1773 CTGCCAGCTGTGTGTATCAGCCAGACACATTGAGCGAGTATGTCCCGCACACCT 1714
|||||
Db 421 tcagttatctg 430
|||||
Cp 1713 TCAGTATCTG 1704
|||||

RESULT 8

LOCUS N23118 430 bp mRNA EST 28-DEC-1995
DEFINITION yx67cl2.s1 Homo sapiens cDNA clone 266806 3' similar to SP:A41234
DEFINITION A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;
ACCESSION N23118
NID gi137268
KEYWORDS EST.
SOURCE human clone=266806 primer=ml3 -40 forward library=Soares melanocyte
2NbHM vector=pT7T3D (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel-Not I Reitel-Eco RI Male.
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGCGCGCGAGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 430)

REFERENCE

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaaks,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.

TITLE
JOURNAL
COMMENT

Contact: Wilson RK

WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 288

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1137268

FEATURES

source

Location/Qualifiers
1..430
/organism="Homo sapiens"
/clone="266806"
/note="human"

BASE COUNT 114 a 107 c 115 g 93 t 1 others

Query Match 19.3%; Score 420; DB 113; Length 430;

Best Local Similarity 99.3%; Pred. No. 0.00e+00;

Matches 427; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 caggctctgagtatttatttcagtaagtagtctccagggaagactgggggaaatat 60
|||||
Cp 2132 CAGGCTCTGAGTATTTATTTTCAGTTAAAGTAGTCTCCACGGGAAGACTGGGGGAATAT 2073
|||||
Db 61 aggtgtttctgcaactccaggaaaatacacagcatcatatgagtagtactcagacctgctg 120
|||||
Cp 2072 AGGTGTTTCTCAATCCAGGAATAACAGCATCATATGAGTACTCAGACTGCTG 2013
|||||
Db 121 ccaactgaggagggtgtatttccaccaatgggacaaagcagaagatcggggtagacg 180
|||||
Cp 2012 CCCACTGAGGAGGGGCTGTTCTCACCAATGGCAAGAGCAGAGATCGGGGTAGACG 1953
|||||
Db 181 cagccagtgtactgctgtatgtggcaactgggtacggagagtgcttcttcataagtct 240
|||||
Cp 1952 CAGCCAGTGTACTGCTGTATGTGGCACTGGGTAGCGAGAGTCTTGTCATAGTCT 1893
|||||
Db 241 ggcctatatatcagagatgcaaggaccacacagccatcaacacagcaagatgccacgat 300
|||||
Cp 1892 GGCCTATATCAGAGATGCAAGGACCACAGCCATCAACACAGCAGATGCCACGAT 1833
|||||
Db 301 cagcggaaactgccaaagctcttctgaccaggcatgataagctgggtgcttgacca 360
|||||
Cp 1832 CAGCGAACCCTGCCAAGGCTCTTCTTGACAGGCATGATAGTGGGTGCT-GACCA 1774
|||||
Db 361 ctgccaggtntgtatcagccagagacacattgaggcagtatgtcccccagaccct 420
|||||
Cp 1773 CTGCCAGGCTCTTGTTATCAGCAGACACATTGAGGCAGTATGTCCCCGAGCCACCT 1714
|||||
Db 421 tcagtatctg 430
|||||
Cp 1713 TCAGTATCTG 1704

9

LOCUS N23111 431 bp mRNA EST 28-DEC-1995
DEFINITION yx67al2.s1 Homo sapiens cDNA clone 266782 3' similar to SP:A41234
A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ?
ACCESSION N23111
NID g1137261
KEYWORDS
SOURCE

human clone=266782 primer=ml3 -40 forward library=Soares melanocyte
2NBHM vector=pt7T30 (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGAGCGGCGAGTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 431)

Authors Hillier,L., Clark,N., Duboue,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaaks,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 286
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1137261

FEATURES

source

Location/Qualifiers
1..431
/organism="Homo sapiens"
/clone="266782"
/note="human"

BASE COUNT 114 a 105 c 116 g 95 t 1 others

ORIGIN

Query Match 18.6%; Score 405; DB 113; Length 431;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 424; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

Db 1 caggctctgagtatttatttcagtaagtagtctccagggaagactgggggaaatat 60

|||||
Cp 2132 CAGGCTCTGAGTATTTATTTTCAGTTAAAGTAGTCTCCACGGGAAGACTGGGGGAATAT 2073

|||||
Db 61 aggtgtttctgcaactccaggaaaatacacagcatcatatgagtagtactcagacctgctg 120

Cp 2072 AGGTGTTCTGTCAACTCCAGGAAATCACAGCATCATATCAGAGTACTCAGACCTGCTG 2013
|||||
Db 121 cccactgaggggggtattctcaccatgggacaagagagaagatgcggggtagagc 180
|||||
Cp 2012 CCCACTGAGGAGGGGCTGTTCTCACCAGTCCGACAGCAGCAGATCGGGGTAGAGG 1953
|||||
Db 181 caqccagtgcctgctatgtgccaactgggtacggagaagctctgtttcacaagtct 240
|||||
Cp 1952 CAGCCAGTGCAGTCTGCTATGTGGCACTGGGTACGGAGAGTCTTGGTTCAAGTCT 1893
|||||
Db 241 ggcctatatcatcagagatcgaagaccacagccatcaaccagcaagatgccacgat 300
|||||
Cp 1892 GCGCTATATCAGAGATCAAGGACCCAGCCATCAACAGCAGCAAGATGCCACCAT 1833
|||||
Db 301 cagcggaacctgcccaagccttctcttgaccagcagatgaagctgggtgctgacca 360
|||||
Cp 1832 CAGCGGAACCTGCCAAGGCT-GCTTCTTGACCGGCATGATAAGCTGGGTGCTGACCA 1774
|||||
Db 361 ctgcagcgtgtgtgatcagcagagacacacatttgaggcagtgatgtgttcccgagccacc 420
|||||
Cp 1773 CT-GCCAGGCTGTGTATCAGCCAGCAGACACATT-GAGGCGATGTGTCCCGAGCCACC 1716
|||||
Db 421 cttcagcat 429
|||||
Cp 1715 CTTCAAGTAT 1707

RESULT 10

ID HS111256 standard; RNA; EST; 431 BP.
AC N23111;
DT 29-DEC-1995 (Rel. 46, Created)
DT 29-DEC-1995 (Rel. 46, Last updated, Version 1)
DE yx67a12.sl Homo sapiens cDNA clone 266782 3' similar to SP:A1234
DE A1234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR -
KW EST.

OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-431
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevasakis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@washington.wustl.edu High quality sequence stops: 286 Source: IMAGE
CC Consortium, LNL This clone is available royalty-free through LNL
CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further
CC information. NCBI gi: 1137261

FC Key Location/Qualifiers

EH
FT source
FT 1..431
FT /organism="Homo sapiens"
FT /clone="266782"
FT /note="human"
FT mRNA
FT <1..>431
SQ Sequence 431 BP; 114 A; 105 C; 116 G; 95 T; 1 other;

Query Match 18.6%; Score 405; DB 129; Length 431;

Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 424; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

Db 1 caggtctcgagtagtattttcagtttaataagtagtctccccaggaagaactgggggaaatat 60
|||||
Cp 2132 CAGGCTCTGAGTATTTATTTTCAGTTAAATAGTAGTCTCCAGGGAAGACTGGGGAAATAT 2073
|||||
Db 61 aggtgtttctgtcaactccaggaaatcacagcatcatatgagtagtactcagacctgtg 120
|||||
Cp 2072 AGTGTCTTCTGCTCACTCCAGGAAATCACAGCATCATATCAGAGTACTCAGACCTGCTG 2013
|||||
Db 121 cccactgaggggggtctattctcaccatgggacaagacagaaatgcggggtagagc 180
|||||
Cp 2012 CCCACTGAGGAGGGGCTGTTCTCAACCAATGGGCAAGCAGCAGATGCCGGGTAGAGG 1953
|||||
Db 181 cagccagtgactgctgctatgtggcaactgggtacggagaagcttctgttcataagtct 240
|||||
Cp 1952 CAGCCAGTGCAGTCTGCTATGTGGCAACTGGGCTACGGAGAAGCTTGTCTTCAAGTCT 1893
|||||
Db 241 ggcctatatcatcagagatcgaagaccacagccatcaaccagcaagatgccacgat 300
|||||
Cp 1892 GCGCTATATCAGAGATCAAGGACCCAGCCATCAACAGCAGCAGATGCCACCAT 1833
|||||
Db 301 cagcggaacctgcccaagccttctcttgaccagcagatgaagctgggtgctgacca 360
|||||
Cp 1832 CAGCGGAACCTGCCAAGGCT-GCTTCTTGACCGGCATGATAAGCTGGGTGCTGACCA 1774
|||||
Db 361 ctgcagcgtgtgtgatcagcagagacacatttgaggcagtgatgtgttcccgagccacc 420
|||||
Cp 1773 CT-GCCAGGCTGTGTATCAGCCAGCAGACACATT-GAGGCGATGTGTCCCGAGCCACC 1716
|||||
Db 421 cttcagcat 429
|||||
Cp 1715 CTTCAAGTAT 1707

RESULT 11

LOCUS N40290 521 bp mRNA EST 22-JAN-1996
DEFINITION yx80c07.r1 Homo sapiens cDNA clone 268044 5' similar to PIR:A53668
A53668 glycoprotein gp100 precursor, melanocyte lineage - human ;
ACCESSION N40290
NID g1163835
KEYWORDS EST.

SOURCE human clone=268044 primer=T7 library=Soares melanocyte 2NbHM
vector=pf7T3D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male. 1st strand
cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTCAAGTGGGAGCGCGAGTGTGTGTGTGTGTGTGTGT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pf7T3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

AUTHORS
1 (bases 1 to 521)
Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
Trevasakis E., Waterston R., Williamson A., Wohlmann P. and

Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.wustl.edu
High quality sequence stops: 340
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.

NCBI gi: 1163835

Location/Qualifiers

source
1..521
/organism="Homo sapiens"
/clone="268044"
/note="human"

mRNA
BASE COUNT 122 a 150 c 135 g 110 t 4 others
ORIGIN

Query Match 17.5%; Score 380; DB 118; Length 521;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 405; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Db 1 atgacctagtggctatctcgtggaagctgacctctctacacctgggaccttggagaca 60
|||||
Qy 772 ATGACCCAGTGGCTATCTGGCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACA 831

Db 61 gtagtgaacctgatctctcgggcaacttgggtcactcatacttaacctggagcctggcc 120
|||||
Qy 832 GTAGTGGAAACCCCTGATCTCTCGGGCACTTGTGGTCACTCATACTTACCTGGAGCCTGGCC 891

Db 121 cagtcactgcccagtggtcctcaggtgcccattcctctcactcactcctctggtcctccc 180
|||||
Qy 892 CAGTCACTGCCAGCGGTGCTCGAGGCTGCATTCCTCTCACCTCCTGTGGCTCTCTCCC 951

Db 181 cagttcaggccaccagatggcagagcccaactgcagagggccctaacaccacagctg 240
|||||
Qy 952 CAGTTCCAGGCACACAGATGGGCACAGGCCAACCTGCAGAGGCCCTTAACACACAGCTG 1011

Db 241 gccagtgccctactacagaagttgtgggtactacacctggtcagggcccaactgcagagc 300
|||||
Qy 1012 GCCAAGTGCCTACTACAGAACTTGTGGTACTACACTGCTGAGGGCCCACTGCAGACC 1071

Db 301 cctctggaaccacatctgtgcaggtggcaaccactgaaagtataaagcactgcacctgtgc 360
|||||
Qy 1072 CTTCTGGNACCACATCTGTGCAGGTGCCAACCACTGAGTGTATTAAGCATGCACCTGTGC 1131

Db 361 agaagtcaccaactgcagagagcacaggtatggacacctggagaggtgcc 410
|||||
Qy 1132 AGA-TGCC-ANCTGCAGAGAGCACAGGTATG-ACACCTG-AGAAGGTGCC 1177

RESULT 12

ID HS290283 standard; RNA; EST; 521 BP.
AC N40290;
DT 26-JAN-1996 (Rel. 46, Created)
DT 26-JAN-1996 (Rel. 46, Last updated, Version 1)

DE yx80c07.r1 Homo sapiens cDNA clone 268044 5' similar to PIR:A53668
DE A53668 glycoprotein gp100 precursor, melanocyte lineage - human ;
KW EST.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae.

RN [1]

RP 1-521

RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,

RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,

RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;

RT "The WashU-Merck EST Project";

RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University

CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,

CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@wustl.wustl.edu High quality sequence stops: 340 Source: IMAGE

CC Consortium, LNL This clone is available royalty-free through LNL

CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further

CC information. NCBI gi: 1163835

FH Key

FH Location/Qualifiers

FT source

1..521

/organism="Homo sapiens"

/clone="268044"

/note="human"

FT mRNA

<1..>521

SQ Sequence 521 BP; 122 A; 150 C; 135 G; 110 T; 4 other;

Query Match 17.5%; Score 380; DB 131; Length 521;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 405; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Db 1 atgacctagtggctatctcgtggaagctgacctctctacacctgggaccttggagaca 60
|||||
Qy 772 ATGACCCAGTGGCTATCTGGCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACA 831

Db 61 gtagtgaacctgatctcgtggaacttgggtcactcatacttaacctggagcctggcc 120
|||||
Qy 832 GTAGTGGAAACCCCTGATCTCTCGGGCACTTGTGGTCACTCATACTTACCTGGAGCCTGGCC 891

Db 121 cagtcactgcccagtggtcctcaggtgcccattcctctcactcactcctctggtcctccc 180
|||||
Qy 892 CAGTCACTGCCAGGTGCTCTGCAGGCTGCCATTCTCTCACCTCTGTGGCTCTCTCCC 951

Db 181 cagttcaggccaccagatgggcagagcccaactgcagagggccctaacaccacagctg 240
|||||
Qy 952 CAGTTCCAGGCACACAGATGGGCACAGGCCAACCTGCAGAGGCCCTTAACACACAGCTG 1011

Db 241 gccaaagtgcctactacagaagttgtgggtactacacctggtcagggcccaactgcagagc 300
|||||
Qy 1012 GCCAAGTGCCTACTACAGAACTTGTGGGTACTACCTGTGTCAGGGGCCCACTGCAGACC 1071

Db 301 cctctggaaccacatctgtgcaggtggcaaccactgaaagtataaagcactgcacctgtgc 360
|||||
Qy 1072 CTTCTGGNACCACATCTGTGCAGGTGCCAACCACTGAGTGTATTAAGCATGCACCTGTGC 1131

Db 361 agaagtcaccaactgcagagagcacaggtatggacacctggagaggtgcc 410
|||||

Qy 1132 AGA-TGCC-ANCTGCAGAGAGCACAGGTATG-ACACCTG-AGAAGGTGCC 1177

RESULT 13

Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	4698	100.0	661	15	R84854	MART-1 melanoma anti		0.00e+00
2	4684	99.7	661	14	R78616	Melanoma associated a		0.00e+00
3	4679	99.6	661	15	R84855	MART-1 melanoma anti		0.00e+00
4	238	5.1	30	6	R30054	Melanoma cell surface		1.67e-08
5	206	4.4	400	1	P19141	Sequence of preprospa		3.39e-06
6	172	3.7	166	2	R07670	Intestinal mucin dedu		8.21e-04
7	176	3.7	278	2	R07671	Intestinal mucin dedu		4.35e-04
8	174	3.7	278	3	R12535	SMC-41 intestinal mu		5.98e-04
9	169	3.6	907	14	R80144	EBV gp350/220.		1.32e-03
10	169	3.6	907	3	P50073	Epstein-Barr virus (E		1.32e-03
11	165	3.5	166	3	R12534	SMC-40 intestinal mu		2.48e-03
12	162	3.4	2035	11	R57141	Host cell factor prot		3.97e-03

ALIGNMENTS

RESULT	1
ID	R84954 standard; Protein; 661 AA.
AC	R84954;
DT	08-MAY-1996 (revised)
DT	20-APR-1996 (first entry)
DE	MART-1 melanoma antigen cDNA25.
KW	CDNA25; MART-1; melanoma antigen recognised by T-cell;
KW	gp100 antigen derivative; melanoma; metastatic melanoma;
KW	tumour-associated antigen; immunogen; diagnosis; prognosis;
KW	prophylaxis; therapy; vaccine.
OS	Mammalian sp.

Key	Peptide	Location/Qualifiers
FT	/label= antigenic peptide	457..466
FT	/note="see R84199"	
FT	W09529193-A2.	
PD	02-NOV-1995.	
PF	21-APR-1995; U05063.	
PF	22-APR-1994; US-231565.	
PF	05-APR-1995; US-417174.	
PA	(US\$) US SEC DEPT HEALTH.	
PI	Kawakami Y, Rosenberg SA;	
DR	WPI; 95-382963/49.	
DR	N-PSDB; T02716.	
PPT	DNA encoding melanoma antigens reco	
PPT	vaccine, host cells and antibodies,	
PPT	immunise animal against melanoma.	

PS Claim 81; Fig 5A; 184pp; English.
 CC cDNA2 is a melanoma antigen (MART-1) which is recognized by
 CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-
 CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source
 CC of immunogenic peptides (see R84199) which are optionally modified
 CC (see R84200-R84211) to enhance their binding to a MHC molecule and
 CC used in medicaments, especially vaccines, for the treatment or
 CC prevention (by immunisation) of melanoma. Antibodies against cDNA2
 CC and its immunogenic peptides may be used in the detection and
 CC isolation of the antigen from a sample, the detection of which is
 CC indicative of a disease state (melanoma or metastatic melanoma).
 SQ Sequence 661 AA;

Query Match 100.0%; Score 4698; DB 15; Length 661;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mdlvklrcllhavlgallavatkvpnqdlgvsrglrltkawnrqlypewteaqrldc 60
 QY 1 MDLVKRLCLLHLLAVLGALLAVGATKVPNNQDLGVSRLTKAWNRQLYPETEAQRLLDC 60

Db 61 wrqgqvalksndqptliganasfaialnfpgsqkvlpdqgviwvntlingsqvvggqp 120
 QY 61 WRGQVSLKVSNDGPTLIGANASFSIALNFPQSGKVLDPDQGVIVWNTLINGSQVWGGQP 120

Db 121 vvpqetddaciipdggpcpgagswskrsfyvwtgqygwglgpgvsglsgtgramlg 180
 QY 121 VYPQETDDACIIPDGGPCPGAGSWSKRSFYVWTKTWGQYMQFLGGPVSGLSIGTGRAMLG 180

Db 181 thtmvtyvhrgrsryvplahsaeafitdqvpsvseqlraldgnkhflrncpltf 240
 QY 181 THTMVTVVHRGRSRYVPLAHSAEAFITDQVPFSVSQRLALDGNKHFLRNQPLTF 240

Db 241 alqlhdpsylaeadlsytwdfdsagtlisralvvtthylepgvtaqvlgaaipits 300
 QY 241 ALQLHDPSCYLAEADLSYTWDFGDSGTLISRALVVTHTYLEPVTAAQVVLQAALPITS 300

Db 301 cgsfpvpgttdghrptaeapnttaqvpvteevgttppqaptaeptsgtsqvpttevis 360
 QY 301 CGSFPVPGTTDGHPTAEAPNTTAQVPVTEEVGTTTTPQAPTAEPTSGTSQVPTTEVIS 360

Db 361 tapvqmpataestgmtpekvpvsevmgttlaemstpeatgmtpaevsiivlsgttaaqvt 420
 QY 361 TAPVQMPATAESTGMTPEKVPVSEVMGTTLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420

Db 421 tewettarelpipepgcdassimstesitgslgplldgtatrlrvkrqvpidevlyry 480
 QY 421 TEWETTARELPIPEPGDASSIMSTESITGSLGPLLDGTATRLRVKRQVPIDCVLYRY 480

Db 481 gsfvrtldivqgieaealqavpsgegafeltvcscqglpkacameispgqcpaqr1 540
 QY 481 GSFVRTLDIVQIEAEALQAVPSGEGAFELTVSCQGLPKACAMEISPGQCPAQR1 540

Db 541 cqpvlpsacqlvhlkkggegtcynlvaladtnslavstqlimpqgeaglgqvpliv 600
 QY 541 CQPVLPSACQLVHLKKGSGCTCYNLVALADTNSLAVSTQLIMPQGEAGLGQVPLIV 600

Db 601 qillvlmavvlasliyrllmkqdfsvpqlphesshlrlprifcscplgenspllsqgg 660
 QY 601 QILLVLMVVLASLIYRRLMKQDFSVPLPHSSHLLRLPRIFCSCPLGENSPILSQGG 660

Db 661 V 661
 QY 661 V 661

RESULT 2
 ID R78646 standard; Protein; 661 AA.
 AC R78646;
 DT 22-JAN-1996 (first entry)
 DE Melanoma associated antigen gp100.
 KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
 KW identification; tumour; gp100.
 OS Homo sapiens.
 PN EP-668350-A1.
 PD 23-AUG-1995.
 PF 14-FEB-1995; 200348..
 PR 16-FEB-1994; EP-200337.
 PR 21-DEC-1994; EP-203709.
 PA (AIKU) AKZO NOBEL NV.
 PI Adema GJ, Figdor CG;
 DR WPI; 95-284790/38.
 PT Melanoma associated antigen gp100 - used in vaccines and for the
 PS detection of tumours
 PS Claim 1; Page 22-24; 40pp; English.
 CC Immunogenic peptides derived from the melanoma associated antigen
 CC may be used in the production of vaccines. Nucleotide sequences
 CC encoding the immunogenic peptides may be used as primers and probes
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes
 CC capable of binding to the melanoma associated antigen can be
 CC cultured ex vivo and returned to melanoma particles, and when
 CC radiolabelled, they may be used to identify tumour deposits.
 SQ Sequence 661 AA;

Query Match 99.7%; Score 4684; DB 14; Length 661;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;
 Matches 660; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdlvklrcllhavlgallavatkvpnqdlgvsrglrltkawnrqlypewteaqrldc 60
 QY 1 MDLVKRLCLLHLLAVLGALLAVGATKVPNNQDLGVSRLTKAWNRQLYPETEAQRLLDC 60

Db 61 wrqgqvalksndqptliganasfaialnfpgsqkvlpdqgviwvntlingsqvvggqp 120
 QY 61 WRGQVSLKVSNDGPTLIGANASFSIALNFPQSGKVLDPDQGVIVWNTLINGSQVWGGQP 120

Db 121 vvpqetddaciipdggpcpgagswskrsfyvwtgqygwglgpgvsglsgtgramlg 180
 QY 121 VYPQETDDACIIPDGGPCPGAGSWSKRSFYVWTKTWGQYMQFLGGPVSGLSIGTGRAMLG 180

Db 181 thtmvtyvhrgrsryvplahsaeafitdqvpsvseqlraldgnkhflrncpltf 240
 QY 181 THTMVTVVHRGRSRYVPLAHSAEAFITDQVPFSVSQRLALDGNKHFLRNQPLTF 240

Db 241 alqlhdpsylaeadlsytwdfdsagtlisralvvtthylepgvtaqvlgaaipits 300
 QY 241 ALQLHDPSCYLAEADLSYTWDFGDSGTLISRALVVTHTYLEPVTAAQVVLQAALPITS 300

Db 301 cgsfpvpgttdghrptaeapnttaqvpvteevgttppqaptaeptsgtsqvpttevis 360
 QY 301 CGSFPVPGTTDGHPTAEAPNTTAQVPVTEEVGTTTTPQAPTAEPTSGTSQVPTTEVIS 360

Db 361 tapvqmpataestgmtpekvpvsevmgttlaemstpeatgmtpaevsiivlsgttaaqvt 420
 QY 361 TAPVQMPATAESTGMTPEKVPVSEVMGTTLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420

Db 421 tewettarelpipepgcdassimstesitgslgplldgtatrlrvkrqvpidevlyry 480

Qy 421 TEWVETARELP IPEEGPDASSIMSTESITGSLG PLLDGTATLRLVKRQVP LDCVLYR 480
 Db 481 gsfvltldivqgisaeilqavpsgegdafeltvscggllpkeacmeispgcqpqrll 540
 Qy 481 GSFVTLTDIVQGISAEILQAVPSGEGDAFELTVSCGGLLPKEACMEISSPGCQPQRLL 540
 Db 541 cqpvlspacqlvlhqlkgsgegtcylnvaladtnslavstqlmpqgeaglgqvpliv 600
 Qy 541 CQPVLSPACQLVLHQLKGSGETCYLNVALADTNSLAVSTQLMPQGEAGLGQVPLIV 600
 Db 601 gillivmavvlasliyrllmkqdfsvqqlphssshwlrllprifcscpiqenpillesqq 660
 Qy 601 GILLIVMAVVLASLIYRRLMKQDFSVQQLPHSSSHWLRLLPRIFCSCPIGENSPILLESQQ 660
 Db 661 v 661
 Qy 661 V 661

RESULT 3
 -ID R84855 standard; Protein; 661 AA.
 AC R84855;
 DT 08-MAY-1996 (revised)
 DT 20-APR-1996 (first entry)
 DE MART-1 melanoma antigen gp100.
 KW gp100; MART-1; melanoma antigen recognised by T-cell;
 KW cDNA25 antigen derivative; melanocyte; melanoma;
 KW metastatic melanoma; tumour-associated antigen; immunogen;
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.
 OS Mammalian sp.

FH Key Location/Qualifiers
 FT Peptide 154..163
 FT /label= C9-154 immunogenic_peptide
 FT Peptide 208..217
 FT /label= C9-209 immunogenic_peptide
 FT /note= "see R84210"
 FT Peptide 280..288
 FT /label= C9-280 immunogenic_peptide
 FT /note= "see R84208"
 FT Peptide 457..266
 FT /label= immunogenic_peptide
 FT Peptide 476..485
 FT /label= immunogenic_peptide
 PN W09529193-AZ.
 PD 02-NOV-1995.
 PE 21-APR-1995; U05063.
 PR 22-APR-1994; US-231565.
 PR 05-APR-1995; US-417174.
 PA (USSH) US SEC DEPT HEALTH.
 PI Kawakami Y, Rosenberg SA;
 DR WPI; 95-382963/49.
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
 PT vectors, host cells and antibodies, used to detect, treat and
 PT immunise animal against melanoma.
 PS Claim 81; Fig 7A; 184pp; English.
 CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which
 CC is recognized by T-lymphocytes, and is a derivative of the
 CC melanoma-specific antigen cDNA25 (see R84854). gp100 is a source
 CC of immunogenic peptides which are optionally modified to enhance
 CC their binding to a MHC molecule, and used in medicaments,
 CC especially vaccines, for the treatment or prevention (by
 CC immunisation) of melanoma. Antibodies against cDNA2 and its
 CC immunogenic peptides may be used in the detection and isolation

CC of the antigen from a sample, the detection of which is indicative
 CC of a disease state (melanoma or metastatic melanoma).
 SQ Sequence 661 AA;

Query Match 99.6%; Score 4679; DB 15; Length 661;
 Best Local Similarity 99.7%; Pred. No. 0.00e+00;
 Matches 659; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 mdlvlkrcellhlavigallavgatkvprnqdlwvsrqlrtkwnrlpewteaqrldc 60
 Qy 1 MDLVLKRCCLLHLAVIGALLAVGATKVPRNQDMLVSRQLRTKWNRLPEWTEAQR LDC 60
 Db 61 wrggqvskvnsdgtlliganasfialnfpqskvlpdgqvlwvnttiingeqvqgqp 120
 Qy 61 WRGGQVSKVNSDGTLLIGANASFIALNFPQSKVLPDGQVLWVNTTIINGSQVWGQP 120
 Db 121 vvpqetddacifpdgpcpsqswskrfvyvkwgtwgywvlgpvgvsgisqvgvramlg 180
 Qy 121 VYPQETDDACIFPDGPCPSQSWSKRFVYVKWGTWGYWV LGPVGVSGLSIGTGRAMLG 180
 Db 181 thtmvttvhhrrgsrsvplahssafitdqvfsvsveqlraldgnkhflrnqpltf 240
 Qy 181 THTMVTVHHRGRSRSVPLAHSSAFITDQVPESVSVQLRALDGNKHFRLNQPLTF 240
 Db 241 alqlhdpssylaeadlsytwdfgdsotllisralvvthtylepqpvtavqlqaapls 300
 Qy 241 ALQLHDPSSYLAEADLSYTWDFGDSOTLLISRALVVTHTYLEPQPVTAVVQAALPLTS 300
 Db 301 cgspevpgttddhrrtaeapntaagvpttevgttgcgaetaesgtsqvpttevis 360
 Qy 301 CGSPEVPGTTDGHRTAEAPNTAGQVPTTEVVGTTCQAPTAEPSGTSVQVPTTEVIS 360
 Db 361 tapvqmpataestgmtpekvpvsevmgttlaemstpeatgmtpaevsvivlsgttaqvtt 420
 Qy 361 TAPVQMPATAESTGMTPEKVPSEVMGTTLAEMSTPEATGMTPAEVSIVLSGTTAAQVTT 420
 Db 421 tewvettarelpipeegpdassimstesitgslgplldgtatrlrvkrqvpldcvlyr 480
 Qy 421 TEWVETARELP IPEEGPDASSIMSTESITGSLGPLLDGTATRLRVKRQVP LDCVLYR 480
 Db 481 gsfvltldivqgisaeilqavpsgegdafeltvscggllpkeacmeispgcqpqrll 540
 Qy 481 GSFVTLTDIVQGISAEILQAVPSGEGDAFELTVSCGGLLPKEACMEISSPGCQPQRLL 540
 Db 541 cqpvlspacqlvlhqlkgsgegtcylnvaladtnslavstqlmpqgeaglgqvpliv 600
 Qy 541 CQPVLSPACQLVLHQLKGSGETCYLNVALADTNSLAVSTQLMPQGEAGLGQVPLIV 600
 Db 601 gillivmavvlasliyrllmkqdfsvqqlphssshwlrllprifcscpiqenpillesqq 660
 Qy 601 GILLIVMAVVLASLIYRRLMKQDFSVQQLPHSSSHWLRLLPRIFCSCPIGENSPILLESQQ 660
 Db 661 v 661
 Qy 661 V 661

RESULT 4
 ID R30054 standard; peptide; 30 AA.
 AC R30054;
 DT 05-MAY-1993 (first entry)
 DE Melanoma cell surface protein N-terminal sequence.
 KW N-terminal; membrane; bound; ME20; antigen; human; melanoma; tumour;
 KW antibody; immune response.

Future research

CC mucin prodn. suc

```
Query Match      3.7%; Score 174; DB 3; Length 278;
Best Local Similarity 28.0%; Pred. No. 5.98e-04;
Matches 46; Conservative 40; Mismatches 70; Indels 8; Gaps 6;

Db      6 aptpistttttvptptgtgtptttttttttttvptptgt---qtpvtvlltttt 62
      :| | : : | | : | | | : | | : | | : | | | | | | | | | | | |
QY      304 SPVCTTGDGHRPTAENNTAGQVPTTEVGTTPGQAPTAEPSTGTSVQVPTTEVISTAP 363
      :| | : : | | : | | | : | | : | | : | | | | | | | | | | | |

Db      63 tmtptptstktvt-pittttttvairptgt-qtpmipistttttvptptpigs 120
      || : : | : | : | | : | | : | | : | | : | | : | | : | |
QY      364 VQMPAESTGMPKVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEW 423
      | : | | | : : : : | | : : : | | : : : | | : : : | |

Db      121 tgpptstapiaelttntppesetpqtarstsepltestill 164
      | : | | | : : : : | | : : : | | : : : | | : : : | |
QY      424 VET-T-ARELPPEPEGDASSIMST-ESITGSLGCLLDGTATL 464

RESULT 9
ID      R80144 standard; Protein; 907 AA.
AC      R80144;
DT      18-JAN-1996 (first entry)
DE      EBV gp350/220.
KW      EBV; gp350; gp220; gp350/gp220; non-splicing variant; vaccine.
OS      Epstein-Barr virus.
FH      Key Location/Qualifiers
FT      Peptide 1..18
FT      /label= Sig_peptide
FT      Region 861..881
FT      /note= "transmembrane region"
PN      WO952848-A1.
PD      26-OCT-1995.
PF      13-APR-1995; U04611.
PR      18-APR-1994; US-229791.
PA      (AVIR-) AVIRON.
PI      Jackson WT, Spaete R;
DR      N-PSDB; T04821.
DR      WPI; 95-373802/48.
PT      New DNA encoding a homogeneous gp350 protein - can be used for
PT      preventing and treating Epstein-Barr virus-related diseases or
PT      conditions
PS      Disclosure; Fig.1; 6lpp; English.
CC      The donor and acceptor splice sites of the EBV gene encoding gp350/
CC      220 are mutated by replacement of native nucleotides by non-native
CC      nucleotides, without altering the encoded amino acid sequence,
CC      resulting in elimination of gp220 prodn. Recombinant homogeneous
CC      gp350, useful in vaccines, is expressed in mammalian or insect cell
CC      hosts.
SQ      Sequence 907 AA;

Query Match      3.6%; Score 169; DB 14; Length 907;
Best Local Similarity 23.2%; Pred. No. 1.32e-03;
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;

Db      421 fakaesttsptlntgfdapntttgtpsthvptnltpaetgptvstadvtsptpag 480
      | : : : : | : : | : : | : : | : : | : : | : : | : : | : :
QY      262 FGDSSGTLISRALVTHYLEPQVTA-QVVLQAAIPLTSCGSS-PVPCTDGHRTAE 319
      :| | : : | | : | | | : | | : | | : | | | | | | | | | | | |

Db      481 ttgasvptpspdwngtsgkapdm-tsstspvttptnatgtpavttptnatptp 539
      :| | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY      320 PNNTAGQV-PTTEV-VGTTGQAPTAEPSTGTSVQVPTTEVISTAP-VQMPAESTG 376
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |

Db      540 avttptnatgptlqktsptsvttptnatgptlqktsptsvttptnatgptlqkts 599
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY      377 E-KVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEWETARELP 435
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |

Query Match      3.6%; Score 166 AA.
Best Local Similarity 23.2%; Pred. No. 1.32e-03;
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;

Db      421 fakaesttsptlntgfdapntttgtpsthvptnltpaetgptvstadvtsptpag 480
      | : : : : | : : | : : | : : | : : | : : | : : | : : | : :
QY      262 FGDSSGTLISRALVTHYLEPQVTA-QVVLQAAIPLTSCGSS-PVPCTDGHRTAE 319
      :| | : : | | : | | | : | | : | | : | | | | | | | | | | | |

Db      481 ttgasvptpspdwngtsgkapdm-tsstspvttptnatgtpavttptnatptp 539
      :| | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY      320 PNNTAGQV-PTTEV-VGTTGQAPTAEPSTGTSVQVPTTEVISTAP-VQMPAESTG 376
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |

Db      540 avttptnatgptlqktsptsvttptnatgptlqktsptsvttptnatgptlqkts 599
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY      377 E-KVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEWETARELP 435
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |
```

```
Db      600 pteavttptnatgptvgetspqanathl 630
      | : : : : | : : | : : | : : | : : | : : | : : | : : | : :
QY      436 PEGPDASSI-MSTESITGSLGPLLDGTA-TL 464

RESULT 10
ID      P50073 standard; Protein; 907 AA.
AC      P50073;
DT      17-OCT-1991 (first entry)
DE      Epstein-Barr virus (EBV) outer surface protein.
KW      Epstein-Barr virus; antigen; vaccine.
OS      Epstein-Barr virus.
PN      EP-151079-A.
PD      07-AUG-1985.
PF      28-JAN-1985; 400141.
PR      30-JAN-1984; US-575352.
PR      23-JUL-1984; US-633558.
PA      (UYCH-) UNIV OF CHICAGO.
PI      Kleff E, Tanner J, Hummel M, Belssel C.
DR      WPI; 85-191978/32.
DR      N-PSDB; N50114.
PT      New fragment of Epstein-Barr Virus DNA - useful in vector to
PT      express polypeptide for use in prepn. of vaccine against the
PT      virus and for use in diagnosis.
PS      Claim 2; Page 23-25; 26pp; English.
CC      The sequence encodes an outer surface viral protein of EBV, used
CC      to generate antibodies reacting with the surface proteins of
CC      EBV-infected cells, and in the preparation of a vaccine against EBV.
SQ      Sequence 907 AA;

Query Match      3.6%; Score 169; DB 3; Length 907;
Best Local Similarity 23.2%; Pred. No. 1.32e-03;
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;

Db      421 fakaesttsptlntgfdapntttgtpsthvptnltpaetgptvstadvtsptpag 480
      | : : : : | : : | : : | : : | : : | : : | : : | : : | : :
QY      262 FGDSSGTLISRALVTHYLEPQVTA-QVVLQAAIPLTSCGSS-PVPCTDGHRTAE 319
      :| | : : | | : | | | : | | : | | : | | | | | | | | | | | |

Db      481 ttgasvptpspdwngtsgkapdm-tsstspvttptnatgtpavttptnatptp 539
      :| | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY      320 PNNTAGQV-PTTEV-VGTTGQAPTAEPSTGTSVQVPTTEVISTAP-VQMPAESTG 376
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |

Db      540 avttptnatgptlqktsptsvttptnatgptlqktsptsvttptnatgptlqkts 599
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY      377 E-KVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEWETARELP 435
      | : : | | : | | : | | : | | : | | : | | : | | : | | : | |

Db      600 pteavttptnatgptvgetspqanathl 630
      | : : : : | : : | : : | : : | : : | : : | : : | : : | : :
QY      436 PEGPDASSI-MSTESITGSLGPLLDGTA-TL 464

RESULT 11
ID      R12534 standard; Protein; 166 AA.
AC      R12534;
DT      02-SEP-1991 (first entry)
DE      SMUC-40 intestinal mucin.
KW      SMUC; intestinal mucin; diagnosis; cancer.
OS      Homo sapiens.
FH      Key Location/Qualifiers
FT      Region 3..25
FT      /label= repeat_1
FT      Region 26..48
FT      /label= repeat_2
```


PT Infections
PS Disclosure; Page 39; 71pp; English.
CC HCF is required for the transcription of a number of
CC viral genes, such as the immediate early herpes simplex virus-1
CC genes. Epitopes of the encoded protein can be used in
CC defining functional domains of HCF, identifying compounds that
CC associate with HCF or designing compounds capable of modifying HCF
CC transcription. Such agents can be used to treat viral infections.
SQ Sequence 2035 AA;

Query Match 3.4%; Score 162; DB 11; Length 2035;
Best Local Similarity 26.9%; Pred. No. 3.97e-03;
Matches 57; Conservative 53; Mismatches 85; Indels 17; Gaps 16;

Db 376 ntsslevsgavatadysllqlq-kydipataata-teptpnvpsvp-annpkspapaa 432
Qy 265 SSGTL-ISRALVVT-HTYLEPGVPTAQVWLQAAIPLTSCGSPVPTTDCGHRPTAEAPNT 322

Db 433 aapavqklqvtitllqqaapapptittiqvlt-v-pgasievptaartgvpavlkvt 490
Qy 323 TAGQVPTTEWCTTPGQAPTAEPSTGTSVQVPTTEVISTAPVQMP-TAEST-GMTPE-KVP 380

Db 491 gpqattgtlvtmrpaagkap--vtvslp-agrvmvvvtgsaggtvlg-espqmegm 546
Qy 381 VSEVM-GTTLAEMSTPEATCMTPAEVSIVLSGTTAAQVTTTEWETTARELPPEPEG- 438

Db 547 aalaaaaatqkippseaptvlsvpagttivk 578
Qy 439 PD-ASSIMSTESITGSLGP-LLDGTATLRIVK 468

RESULT 13

ID R53915 standard; Protein; 252 AA.
AC R53915;
DT 07-DEC-1994 (first entry)
KW Polyhydroxyalkanoate decomposing activity polypeptide.
OS Zoogloea ramigera.
PN J06086681-A.
PD 29-MAR-1994.
PF 07-SEP-1992; 279099.
PR 07-SEP-1992; JP-279099.
PA (ELED) DENKI KAGAKU KOGYO KK.
PA (UYKA-) GH KANAGAWA DAIGAKU.
DR (CHIK-) ZH CHIKYU KANKYO SANGYO GIJITSU KENKYU.
DR WP1: 94-140539/17.
DR N-PSDB; 663879.
PT DNA encoding poly-hydroxy-alkanoate decomposing enzyme - useful
PT for preparing poly-hydroxy-alkanoate in large quantities
PS Claim 3; Page 10-11; 15pp; Japanese.
CC R53915 is encoded by ORF 1 of Q63789, which produces a polypeptide
CC which shows polyhydroxyalkanoate decomposing activity. The method
CC used allows the polypeptide to be prepared in large quantities.
SQ Sequence 252 AA;

Query Match 3.3%; Score 157; DB 10; Length 252;
Best Local Similarity 26.0%; Pred. No. 8.66e-03;
Matches 38; Conservative 43; Mismatches 58; Indels 7; Gaps 7;

Db 46 amtvaavtvaavvtnaa-aattvvaatmaavaatttv-aattavv-ttmavaattv 102
Qy 285 PVTAQVWLQAAIPLTSCGSPVPTTDCGHRPTAEAPNTTAGQVPTTEWVGTTCQAPTAE 344

Db 103 vaatmaavv-ttvvaatmaavaattvvaatmaavaattvvaatmaatmvmaettat-maat 160

Qy 345 PSQTTSVQVPTTEVISTAPVQMP-TAESTGTMPEKVPV-SEVMGTTLAEMSTPEATCMTPA 403
Db 161 matmaatatltnaaraamaaraettt 186
Qy 404 EVSIVVLSG-TTAAQVTTTEWETTA 428

RESULT 14
ID R85290 standard; Protein; 671 AA.
AC R85290;
DT 02-APR-1996 (first entry)
DE Streptococcus faecalis autolysin.
KW Lysin; autolysin; culture; lactic acid bacteria; fermentation;
KW cheese; foodstuffs; induction.
OS Streptococcus faecalis.
PN M09531561-A1.
PD 23-NOV-1995.
PF 12-MAY-1995; NL0170.
PR 12-MAY-1994; EP-201353.
PA (UNIL) QUEST INT BV.
PI Buist G, Kok J, Ledebroer AM, Venema G;
DR WP1: 96-010946/01.
PT Lysis of a culture of lactic acid bacteria in, e.g. cheese
PT production - by in situ prodn. of an auto-lysin, regulated by an
PT inducible promoter.
PS Disclosure; Page 66-69; 103pp; English.
CC In situ production of a homologous autolysin or a heterologous
CC autolysin from a food grade Gram positive bacteria, can be used in
CC a process for the lysis of a culture of lactic acid bacteria. The
CC process can be used in the manufacture of products containing
CC cultures of lactic acid bacteria e.g. cheese, where the culture is
CC lysed following the completion of fermentation. The enhanced
CC induction of the autolysin is performed some hours after the
CC fermentation is finished. No extra lysin needs to be added and the
CC lysin does not need to be isolated or encapsulated. The time of
CC lysis can be precisely controlled. This is the Streptococcus
CC faecalis autolysin.
SQ Sequence 671 AA;

Query Match 3.2%; Score 148; DB 15; Length 671;
Best Local Similarity 28.2%; Pred. No. 3.48e-02;
Matches 55; Conservative 49; Mismatches 76; Indels 15; Gaps 13;

Db 27 sttlfsealivssvgtvllpvtataeteeqptn-aevagapgtetgivetpttpttgi 85
Qy 266 SGTLSRALVVTHT--TYLEPGVPTAQVWLQAAIPLTSCGSPVPTTDCGHRPTAEAPNTT 323

Db 86 teq-pttdssttsttstsskettpttsteqptvd--ettvpsqtdss-va-eitpva 140
Qy 324 AGQVPTTEWVGTTCQAPTAE-EPSCGTSVQVPTTEVISTAPVQMP-TAESTGTMPEKVPVS 382

Db 141 ptttseaaavtpdpdevkvparsaaqtfaaleptq-spsefaelarcacpaaqand 199
Qy 383 -EVMGTTLAEMSTPEATCMT-PAEVSIV-VLSGTTAAQVTTTEWETTAR-ELPIPEEG 438

Db 200 lyas-vmaqaalves 213
Qy 439 PDASSIMSTESITGS 453

RESULT 15
ID R75107 standard; Protein; 722 AA.
AC R75107;

Aug 27 08:01

US-08-231-565A-27.rag

17

DT 01-FEB-1996 (first entry)
 DE Bile salt-activated lipase with 16 C-terminal 11-mer repeats.
 KW Bile salt activated lipase; BAL; C-terminal repeat.
 OS Mammalia.
 FH Key Location/Qualifiers
 FT Region 539..722
 FT /note= "contains 16 copies of imperfect 11mer
 FT repeats; the region is made up from a
 FT combination of 7 different sequences"
 PN J07111891-A.
 PD 02-MAY-1995.
 PF 30-SEP-1993; 245079.
 PR 30-SEP-1993; JP-245079.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 DR WPI; 95-196318/26.
 DR N-PSDB; Q90579.
 PT New recombinant bile salt-activated lipase(s), DNA encoding them
 PT and vectors - useful for high level expression of the lipase(s) by
 PT fermenter-culturing.
 PS Disclosure; Page 10-13; 31pp; Japanese.
 CC New lipases are ones in which 1-15 repeats of any of the sequences
 CC R75099-R75105 (encoded by Q90570-Q90576) are attached to the C-
 CC terminus of the bile salt-activated lipase (BAL) of sequence R75098
 CC (encoded by Q90569) and/or in which the amino acid sequence SMTGS
 CC (encoded by Q90578) is attached to the N-terminus of the BAL. The
 CC recombinant lipases can be produced by fermentation in Pichia
 CC pastoris GS 115 transformed host cells. The present sequence
 CC is a protein with 16 copies of the undecapeptide imperfect
 CC repeats at the BAL C-terminus.
 SQ Sequence 722 AA;

Query Match 3.1%; Score 147; DB 14; Length 722;
 Best Local Similarity 24.6%; Pred. No. 4.05e-02;
 Matches 42; Conservative 50; Mismatches 67; Indels 12; Gaps 12;
 Db 547 eatpvptgdsapvpptgdsagppvptgdsagppvptgdsagppv 606
 Qy 282 EPGVTAQVVLQAA-IPLTS-CGSSPVPGTTD-GHRPTAENPTTAGQVPTTEVGTTPG 338
 Db 607 -pvtgd-sgappvp-ptgds-gappvp-ptgds-g-appvpptgdagppvpptgdsagp 659
 Qy 339 QAPTAEFGSTTSVQVPTTEVISTAPVQMPTEASTCMTPEKVPVSEVMGTTLAEMSTPEAT 398
 Db 660 pvpptgdsagppvptgdsapvpptgdsag-appvpptgdsagappvpptd 709
 Qy 399 GMTPA-EVSIVVLSGTTAAQVTTTEWVETTTARELP IPEEGPDASSIMSTE 448

Search completed: Tue Aug 27 08:14:12 1996
 Job time : 78 secs.


```
##molecule type mRNA
##residues 1-161, F', 163-661 ##label KGW
##note nucleotide sequence not given
KEYWORDS glycoprotein
SUMMARY #length 661 #molecular-weight 70255 #checksum 5487

Query Match 99.7%; Score 4684; DB 11; Length 661;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 660; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdlvklrclhlhavigallavgatkvprnqdlgvsrqlrtkawnrqlypewteaqrldc 60
Qy 1 MDLVKRCLLHLAVIGALLAVGATKVP RNQDLGVS RQLRTKAWNRQLYP EWTEAQR LDC 60
Db 61 wrggqvskvndgptlliganasfsialnfpqskvlpdgqvlwvntlingsqvwgqp 120
Qy 61 WRGGQVSKVNDGPTLLIGANASFSIALNFPQSKVLPDGQVLWVNTLINGSQVWGQP 120
Db 121 vvpqetddacifpdgpcpsgswskrfvyvktwgqyvgvlgpvgvsglsgtgramlg 180
Qy 121 VVPQETDDACIFPDGPCPSGSWSKRSFYVVKTWGQYVGLGPGVSGLSIGTGRAMLG 180
Db 181 thtmvtyvhrgrsyrsvplahssafitcdqvpfsvsveqlraldgnkhflrnqpltf 240
Qy 181 THTMVTVVHRGRSRYVPLAHSSAFITDQVPFSVSVEQLRALDGNKHFLRNQPLTF 240
Db 241 alqlhdpsylaeadlsytwdfgdsstllsralvvthtylepqtavqvlqaalplts 300
Qy 241 ALQLHDPSYLAEADLSYTWDFGDSSTLLSRA LVVTHTYLEPQTAVQVLQAALPLTS 300
Db 301 cgssvpvgttddghrptaeapnttagvpttevgttqgaptaeapsgttsvqpttevis 360
Qy 301 CGSSVPVGTDDGHRPTAEAPNTTAGQVPTTEVVGTPQAPTAEPSGTSVQPTTEVIS 360
Db 361 tapvqmpataestgmtpekvpvsevmgtllaemstpeatgmtpaevsiivlsgttaaqvtt 420
Qy 361 TAPVQMPATAESTGMTPEKVPVSEVMGTLLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420
Db 421 tewvettarelpipepegdaasimstesitgslpdlldgtatrlrvkrqpldevlry 480
Qy 421 TEWVETTARELP IPEPEGDASSIMSTESITGSLPDL DGTATRLRVKRQVP LDCVLYRY 480
Db 481 gsfvltldivqgiesaelqavpsgegdafeltvscggglpkacmeisepgcqppagrl 540
Qy 481 GSFVTLTDIVQGESAEILQAVPSEGDAFELTVSCGGGLPKACMEISSPGCQPPAQR L 540
Db 541 cqpvlspacqlvlhqlkgsqgtcylnveladtnslavvstqlimpqagqlgvpliv 600
Qy 541 CQPVLSPACQLVLHQLKGSQGT CYLNV ELADTNSLAVVSTQLIMPQAGQLGVPLIV 600
Db 601 gillvlmavlaaslyrrrlmkqdfsypqlphssahwlrprifscpigenellsgqq 660
Qy 601 GILLVLMAVVLAASLYRRRLMKQDFSYPQLPHSSAHWLRPRIFSCPIGENELLSGQQ 660
Db 661 v 661
Qy 661 V 661
```

```
RESULT 2
ENTRY A41234 #type complete
TITLE melanocyte-specific protein Pmel-17 precursor - human
ORGANISM #formal name Homo sapiens #common name man
DATE 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
```

```
30-Sep-1993
ACCESSIONS A41234
REFERENCE A41234
#authors Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
Kobayashi, Y.; Kim, K.K.
#journal Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title A melanocyte-specific gene, Pmel 17, maps near the silver
coat color locus on mouse chromosome 10 and is in a
syntenic region on human chromosome 12.
#cross-references MIMD:92021023
#accession A41234
##status preliminary
##molecule type mRNA
##residues 1-668 ##label KWO
##cross-references GB:M77348
SUMMARY #length 668 #molecular-weight 70932 #checksum 6409

Query Match 97.8%; Score 4595; DB 11; Length 668;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 659; Conservative 0; Mismatches 2; Indels 7; Gaps 1;

Db 1 mdlvklrclhlhavigallavgatkvprnqdlgvsrqlrtkawnrqlypewteaqrldc 60
Qy 1 MDLVKRCLLHLAVIGALLAVGATKVP RNQDLGVS RQLRTKAWNRQLYP EWTEAQR LDC 60
Db 61 wrggqvskvndgptlliganasfsialnfpqskvlpdgqvlwvntlingsqvwgqp 120
Qy 61 WRGGQVSKVNDGPTLLIGANASFSIALNFPQSKVLPDGQVLWVNTLINGSQVWGQP 120
Db 121 vvpqetddacifpdgpcpsgswskrfvyvktwgqyvgvlgpvgvsglsgtgramlg 180
Qy 121 VVPQETDDACIFPDGPCPSGSWSKRSFYVVKTWGQYVGLGPGVSGLSIGTGRAMLG 180
Db 181 thtmvtyvhrgrsyrsvplahssafitcdqvpfsvsveqlraldgnkhflrnqpltf 240
Qy 181 THTMVTVVHRGRSRYVPLAHSSAFITDQVPFSVSVEQLRALDGNKHFLRNQPLTF 240
Db 241 alqlhdpsylaeadlsytwdfgdsstllsralvvthtylepqtavqvlqaalplts 300
Qy 241 ALQLHDPSYLAEADLSYTWDFGDSSTLLSRA LVVTHTYLEPQTAVQVLQAALPLTS 300
Db 301 cgssvpvgttddghrptaeapnttagvpttevgttqgaptaeapsgttsvqpttevis 360
Qy 301 CGSSVPVGTDDGHRPTAEAPNTTAGQVPTTEVVGTPQAPTAEPSGTSVQPTTEVIS 360
Db 361 tapvqmpataestgmtpekvpvsevmgtllaemstpeatgmtpaevsiivlsgttaaqvtt 420
Qy 361 TAPVQMPATAESTGMTPEKVPVSEVMGTLLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420
Db 421 tewvettarelpipepegdaasimstesitgslpdlldgtatrlrvkrqpldevlry 480
Qy 421 TEWVETTARELP IPEPEGDASSIMSTESITGSLPDL DGTATRLRVKRQVP LDCVLYRY 480
Db 481 gsfvltldivqgiesaelqavpsgegdafeltvscggglpkacmeisepgcqppagrl 540
Qy 481 GSFVTLTDIVQGESAEILQAVPSEGDAFELTVSCGGGLPKACMEISSPGCQPPAQR L 540
Db 541 cqpvlspacqlvlhqlkgsqgtcylnveladtnslavvstqlimpqagqlgvpliv 600
Qy 541 CQPVLSPACQLVLHQLKGSQGT CYLNV ELADTNSLAVVSTQLIMPQAGQLGVPLIV 600
Db 601 gqvplivgillvmaavlaaslyrrrlmkqdfsypqlphssahwlrprifscpigen 660
Qy 601 GQVPLIVGILLVMAAVVLAASLYRRRLMKQDFSYPQLPHSSAHWLRPRIFSCPIGEN 660
```


QY	458	LDGTATLRLWKRQVPLDCLVRYGSFVTLDIVOGTIESAIIQAVPSEGDAFEUWSCQ	517
Db	346	gglpteacmdispggcqlpaqrloqvpypspacqlvhwvilkqsgctcylwsladana1	405
QY	518	GGLPKEACMETSPGCQPPAQLCQVLPSPACQLVHLQILKGGSGCTCYNWSTADTNSL	577
Db	406	amvstqlvmpqceagrlqaplfvgillvltalliaslyrrrrlmkqgsevpjqlphgrrt	465
QY	578	AVVSTQLTMEQCEAGLGQVPLVIGILLVMAVWLASLIYRRRLMKQ--DFSVPQLPHSSS	635
Db	466	qwlrlpwvfracpiqesckplleggvv	491
QY	636	HWLRLPRIFGSCPIGENSPLLSGQV	661

```

5
RESULT#
ENTRY  calphotin      #type complete
TITLE  calphotin - Drosophila
ORGANISM #formal_name Drosophila
DATE    21-Sep-1993 #sequence_revision 21-Sep-1993 #text_change
        21-Sep-1993
ACCESSIONS A47283
REFERENCE  A47283
AUTHORS   Ballinger, D.G.; Xue, N.; Harshman, K.D.
#journal  Proc. Natl. Acad. Sci. U.S.A. (1993) 90:1536-1540
#title    A Drosophila photoreceptor cell-specific protein, calphotin,
        binds calcium and contains a leucine zipper.
#cross-references MUID:93165730
#contents photoreceptor cells
#accession A47283
#status preliminary
#molecule_type nucleic acid
#residues 1-873 ##label BAL
#cross-references NCBI:124958; NCBI:124959
#note     sequence extracted from NCBI backbone
#length 873 #molecular_weight 85445 #checksum 247
SUMMARY

```

Query Match	4.6%	Score 217;	DB 14;	Length 873;
Best Local Similarity	24.2%;	Pred. No.	8.0le-11;	
Matches	48;	Conservative	63;	Mismatches 78; Indels 9; Gaps 8;
Db	122	vappvaactpvapipvaapvi--atppvaaasaptpaavtpvvsviatppvpanttpv	179	
	:	: :	: : : :	: :: : : :
	:	: :	: : : :	: :: : : :
Qy	270	ISRALVWTHLYTEPGVTAAQVLQAALPTSGSGSPVP GTDCHRP TAEAPNTTAQ-QV	328	
Db	180	vaapvaaavaapvpvaplapavapavpvaetpappvpaeipvatipccvaplipavs	239	
	:	: : :	: : : :	: : : :
	:	: : :	: : : :	: : : :
Qy	329	TTEWGTTPGQAPTAEPTSGTTSQVPTTEVI STAPVQMPTAE-STGMTPEKV-P-VSEVM	385	
Db	240	vratkplaaepvpvappatetpvvpaaasaphsvapavetaavva-pvaeateppvaav	298	
	:	: :::: : :	: : :::: : :	: : : : :
	:	: :::: : :	: : :::: : :	: : : : :
Qy	386	GTTLAEMSTPEATGMTPAEVSI-VVLISGTTAAQVTTTEWVETTRALIP-EGPDASS	443	
Db	299	tlttpetpalapvaes	316	
	:	: : : :	: : : :	: : : :
Qy	444	IMSTESITCSGLPILDCT	461	

RESULT	6
ENTRY	A37232
TITLE	mucin, tracheal (AMN-22) - human (fragment)
ORGANISM	#formal_name Homo sapiens #common_name man
DATE	14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 30-Sep-1993

```

ACCESSIONS      A37232
REFERENCE
#authors        Gerard, C.; Eddy Jr., R.L.; Shows, T.B.
#journal         J. Clin. Invest. (1990) 86:1971-1977
#title           The core polypeptide of cystic fibrosis tracheal mucin
                  contains a tandem repeat structure. Evidence for a common
                  mucin in airway and gastrointestinal tissue.
#cross-references M01D:91072667
#accession      A37232
#status          preliminary
#molecule_type mRNA
#residues        1-294 ##label GER
#cross-references GB:M57417
SUMMARY          #length 294 #checksum 1596

```

Query Match	4.5%;	Score 212;	DB 11;	Length 294;	
Best Local Similarity	26.5%;	Pred.No. 2.97e-10;			
Matches	50;	Conservative 44;	Mismatches 89;	Indels 6;	
				Gaps 4;	
Db	16	ptttttttmvtptptltstqtptptp	ptttttttvtptptctctgtcttctattttttvtptptp	75	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	273	ALWVTHYLEPGVTAQVVLQAAIPLTSCGSPVPGTTDGHRTAEAPNTTAQGVPTTEV	332		
Db	76	ittttttvtptptgtptpttp	istttttvtptptgtgt-lltptittttttvtptptp	134	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	333	VGTTFGQAPTAEPSCGT-TSVQVPTTEVISTAPVQMPSTAESTGMTEKVPVSEVMGTILAE	391		
Db	135	tgctgtctattttttvtptptgtptptpt	ittttttttvtptptgtgttctgttptptp	194	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	392	MSPTEATGMTPAEVSIVLISCT-TAAQVTTTTEVWTTTARELPIPEPEG---PDASSIMST	447		
Db	195	ttvtptptp	203		
	:	:	:	:	
	:	:	:	:	
Qy	448	ESITGSLGP	456		

```

RESULT      7
ENTRY
TITLE      spasmolysin precursor - African clawed frog
ALTERNATE_NAMES
ORGANISM   prospasmolysin
DATE       30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change
          02-Jun-1995
ACCESSIONS A28172 #type complete
REFERENCE   A28172 spasmolysin
AUTHORS     Hoffmann, W.
JOURNAL     J. Biol. Chem. (1988) 263:7686-7690
TITLE       A new repetitive protein from Xenopus laevis skin highly
           homologous to pancreatic spasmolytic polypeptide.
CROSS-REFERENCES
MUID:88277968
ACCESSION   A28172
#molecule_type mRNA
#residues 1-400 #label HOF
#cross-references CB:M19971
CLASSIFICATION
KEYWORDS    #superfamily spasmolysin; trefoil homology
           duplication; glycoprotein; skin; tandem repeat
FEATURE
1-20        #domain signal sequence #status predicted #label SIG\
21-400      #product spasmolysin #status predicted #label MAR\
21-69       #product spasmolysin I #status predicted #label SPI\
23-63       #domain trefoil homology #label TRF1\
72-348      #product spasmolysin-glycopeptide #status predicted
           #label SPG\
74-114      #domain trefoil homology #label TRF2\

```



```

##status preliminary
##molecule_type mRNA
##residues 1-322 ##label ALB
##cross-references GB:U03407
KEYWORDS glycoprotein; polymorphism; tandem repeat
FEATURE
1-22
23-322
SUMMARY
    domain signal sequence #status predicted #label SIG\
    #product apomucin #status predicted #label NAT
    #length 322 #molecular-weight 34975 #checksum 1086

Query Match 4.2%; Score 198; DB 12; Length 322;
Best Local Similarity 20.5%; Pred. No. 1.11e-08;
Matches 30; Conservative 48; Mismatches 68; Indels 0; Gaps 0;

Db 118 vtpaqtkpttdsttpaptkkpttdsttpaptkkpttdsttpaptkkpttdsttpaptn 177
    I :: :||::: :: :||::: :| :||::: :: :||::: :: :||::: :: :||
Qy 306 VPCTGDCRRPFAAPNTACQVPTEVVGTTGCPQAPFAEPSGTTSVQVPTTEVISTAPVQ 365
    :||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Db 178 kpttdsttpaptnkpttdsttpaptkkpttdsttpaptkkpttdsttpaptkkpttdstt 237
    :||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Qy 366 MPTAESTGMPEKVPVSEVMGTTIAENSTPEATGMTPAEVSIVLWISGTTAAQVTTTWE 425
    :||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::

Db 238 paptkkpttdsttpaptkkpttdstt 263
    :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: |
Qy 426 TTARELPPIPEGEGDASSIMTESIT 451

RESULT 11
ENTRY S53363 #type fragment
TITLE mucin JER58 - human (fragment)
ORGANISM #formal_name Homo sapiens #common_name man
DATE 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995
ACCESSIONS S53363
REFERENCE S53361
#authors Guyonnet-Duperrat, V.; Audie, J.P.; Debailleul, V.; Laine, A.;
    Buisne, M.P.; Galiegue-Zoutina, S.; Pigny, P.; Degand,
    P.; Aubert, J.P.; Porchet, N.
#journal Biochem. J. (1995) 305:211-219
#title Characterization of the human mucin gene MUC5AC: a consensus
    cysteine-rich domain for lip15 mucin genes?
#accession S53363
##status preliminary
##molecule_type mRNA
##residues 1-279 ##label GUY
##cross-references EMBL:134278
##note the nucleotide sequence is not given in this paper
    #length 279 #checksum 1256
SUMMARY

Query Match 3.9%; Score 183; DB 11; Length 279;
Best Local Similarity 26.9%; Pred. No. 4.91e-07;
Matches 49; Conservative 43; Mismatches 84; Indels 6; Gaps 6;

Db 96 pspvpttst-tsa-pttrttteapissatttatttsggttppspvpttstt-sapttst 152
    :||::: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy 283 PGEVTAQVVLQALPILTSCGSPVPGTTGCHR-TPACAPNTAGQVPTTEVWGTTPGAP 341
    :||::: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 153 tsgp-gttpeapvttt-itcvpttstnsapissttsatttarisgtpptpsvptastts 210
    :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy 342 TAEPGSGTTSVQVPTTEVISTAPVQMPFAESTGMPEKVPVSEVMGTTIAENSTPEATGMT 401
    :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Db 211 asttattggtqtppspvpttatievptttsaatttsaatttsaattggttspvptt 270
    :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy 402 PAEVSIVLWISGTTAAQVTTTWEVTTARELP IPEEGFDASSIMTESITGSLGLDGT 461
    :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

```

Db      271 st 272  
       :  
Qy     462 AT 463  
  
RESULT 12  
ENTRY   A33532          #type fragment  
TITLE   mucin SMOG-40 - human (fragment)  
ORGANISM #formal name Homo sapiens #common name man  
DATE    29-Jan-1990 #sequence_revision 27-Jun-1994 #text_change  
        27-Jun-1994  
ACCESSIONS A33532  
REFERENCE  A33532  
#authors   Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampert,  
            D.T.A.; Kim, Y.S.  
#journal   J. Biol. Chem. (1989) 264:6480-6487  
#title     Molecular cloning of human intestinal mucin cDNAs. Sequence  
            analysis and evidence for genetic polymorphism.  
#cross-references MUID:89197956  
#accession A33532  
#molecule_type mRNA  
#residues_ 1-167 ##label GUM  
#cross-references GB:M22404  
KEYWORDS glycoprotein  
SUMMARY   #length 167 #checksum 7137  
  
Query Match 3.7%; Score 175; DB 11; Length 167;  
Best Local Similarity 27.4%; Pred.No. 3.56e-06;  
Matches 43; Conservative 36; Mismatches 71; Indels 7; Gaps 7;  
  
Db      9 atttmvtpptgtqtltpptittttv-tptptgtgtptstpiattttvtptptpt 67  
       l::l::::|g||l|l:l::|::|::|::|::|::|::|::|::|::|:  
Qy     265 SSGTILSRALVVTH-T-YLEGPVTAQVVLQAIPAITS CGSSPVPGTTDGHRTAEAPNTT 323  
  
Db      68 gtqtpttptittttvpktptgtqt-ptpfp-itttttv-tptptpg-tqtptptpi 123  
       :||||:|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Qy     324 AGQVPTEWVGVTGPAATPEFGSTGSVQPTEVIS TAPVQMPTAESTGMTPEKVPVSE 383  
       ||:::||::|l::|::|::|::|::|::|::|::|::|::|:  
Db      124 tttttvtpptgt-gtptstpitntvtptptpt 159  
       ||:::||::|l::|::|::|::|::|::|::|::|::|::|:  
Qy     384 VMGTLLAESPTPATGMTPAEVSIVVLSTGAQVTT 420  
  
RESULT 13  
ENTRY   S55925          #type complete  
TITLE   probable arabinogalactan protein precursor - tomato  
ORGANISM #formal name lycopersicon esculentum #common name tomato  
DATE    28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change  
        03-Nov-1995  
ACCESSIONS S55925  
REFERENCE  S55925  
#authors   Pogson, B.J.; Davies, C.  
#journal   Plant Mol. Biol. (1995) 28:347-352  
#title     Characterization of a cDNA encoding the protein moiety of a  
            putative arabinogalactan protein from Lycopersicon  
            esculentum.  
#accession S55925  
#status    preliminary  
#molecule_type mRNA  
#residues 1-215 ##label POG  
#cross-references EMBL:Z47980  
SUMMARY   #length 215 #molecular-weight 20812 #checksum 3703
```

```

Query Match      3.7%; Score 172; DB 9; Length 215;
Best Local Similarity 21.8%; Pred. No. 7.43e-06;
Matches 39; Conservative 50; Mismatches 85; Indels 5; Gaps 5;

Db 9 vsilcivvasvtgdtppaaavp=gakagtftppaaatpktpapatapasapptavpva-p 66
   | : :: || :: : :: || | : | | | | | | | | | | | | | | | | | |
Qy 286 VTAQVWIAIPLTISGSSVPVETDGHRTPEANTTAQGVPTTEVGVTTTTCQAPTAEP 345
   | : :: || :: : :: || | : | | | | | | | | | | | | | | | | | |

Db 67 v-tapvtaptipvva-apyasapaspplkapassppqvsppapapevatppavstppaaa 124
   | :: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 346 SGTTSVQVPTTEVISTAPVQMEPTAESTCMTPEKVPVSEVWGTTTAEASTPEATGWTAEV 405

Db 125 pvaapvasettpapapekgtvkgkkgkhaapaspdpdmasppapaseapdpmsdadaa 183
   :: :: :: :: : : : : : : : : : : | | : : | | : : | | : :
Qy 406 STVLSGTTAAQVTTTIEWETTARELP IPEEGPDASSIMST=ESITGSLGLIDGTAT 463

RESULT 14
ENTRY mucin SMUC-41 - human (fragment)
TITLE mucin HAM-1
ALTERNATE_NAMES mucin HAM-1
ORGANISM #Formal name Homo sapiens #common name man
DATE 29-Jan-1990 #sequence_revision 27-Jun-1994 #text_change
      23-Mar-1995

ACCESSIONS B33532; A61257
REFERENCE B33532; A61257
#authors Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampert,
          D.T.A.; Kim, Y.S.
#journal J. Biol. Chem. (1989) 264:6480-6487
#title Molecular cloning of human intestinal mucin cDNAs. Sequence
          analysis and evidence for genetic polymorphism.
#cross-references MUID:89197956
#accession B33532
#molecule_type mRNA
#residues 1-278 ##label GUM
#cross-references GB:M22405
#experimental_source Intestine
REFERENCE A61257
#authors Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.;
          Basbaum, C.B.
#journal J. Clin. Invest. (1991) 87:77-82
#title Human bronchus and intestine express the same mucin gene.
#accession A61257
#status not compared with conceptual translation
#molecule_type mRNA
#residues 'r', 10-33, 'TTS', 37-39 ##label JAN
#experimental_source bronchus
#COMMENT Southern blot analysis suggests that a polymorphic gene encodes
          both intestinal and bronchial mucin.

GENETICS
#gene GDB:MUC2
#map_position 11p15.5
KEYWORDS glycoprotein; polymorphism
SUMMARY #length 278 #checksum 4056

Query Match      3.7%; Score 176; DB 11; Length 278;
Best Local Similarity 28.0%; Pred. No. 2.79e-06;
Matches 46; Conservative 41; Mismatches 69; Indels 8; Gaps 6;

Db 6 aptptpiattttvtptptgtqtppttttttttttttttttttttttttttttttttttttt 62
   : | : : : : | | | | | | | | | | | | | | | | | | | | | | | |
Qy 304 SPVPGTTDGHRTAEANTTAGQVPTTEVGVTTTCQAPTAEPGCTTSVQVPTTISTAP 363

Db 63 tmtotptstktvti-pitttttttttttttttttttttttttttttttttttttttttt 120

```

Qy	364	VQMPTAESTGTMTEKVPVSEVNGTTLAEMSTPEATQMTFAEVSVIWSLSTTAAQVTTTEM	423
Db	121	tgpphtstapiaelttsnppssstqptsrststl1164	
Qy	424	VE-T-ARELIPPEGGPDASSIMST-ESITGSLGLDGTATL	464
RESULT	15		
ENTRY	VOBX1	#type complete	
TITLE	glycoprotein X precursor - equine herpesvirus 1 (strain Ab4p)		
ORGANISM	#formal name equine herpesvirus 1		
#note	host Equus caballus (domestic horse)		
DATE	30-Sep-1992	#sequence_revision 30-Sep-1992	#text_change 17-Feb-1995
ACCESSIONS	H36802		
REFERENCE	A36805		
#authors	Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.		
#submission	submitted to GenBank, March 1992		
#description	The DNA sequence of equine herpesvirus-1.		
#accession	H36802		
#molecule_type	DNA		
#residues	1-797	#label TEL	
#cross-references	GB:M8664		
REFERENCE	A41831		
#authors	Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.		
#journal	Virology (1992) 189:304-316		
#title	The DNA sequence of equine herpesvirus-1.		
#cross-references	MUID:92295566		
#contents	annotation; possible protein-coding frames		
#note	neither amino acid nor nucleotide sequence is given		
GENETICS			
#gene	71		
CLASSIFICATION	#superfamily equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homology glycoprotein; transmembrane protein		
KEYWORDS			
FEATURE			
1-22	#domain signal sequence	#status predicted	#label SIG
23-797	#product glycoprotein X	#status predicted	#label MAT
23-465	#region serine/threonine-rich		
489-797	#domain equine herpesvirus 1 glycoprotein homology		
	#label EHG		
766-790	#domain transmembrane	#status predicted	#label TMN
590	#binding site carbohydrate (Asn (covalent))	#status predicted	
SUMMARY	#length 797	#molecular-weight 80342	#checksum 6547
Query Match	3.6%;	Score 169;	DB 3; Length 797;
Best Local Similarity	17.7%;	Pred. No. 1.54e-05;	
Matches	32;	Conservative 64;	Mismatches 84; Indels 1; Gaps 1;
Db	51	aspttsppttssppsttstspastatgssataatasaapaatasttstetttt	110
Qy	283	PCPVTAQWVLAQAIPLTSCGSPVPGTTDGHRTPEAPNTTACQVPTPEVGTTPGQAPT	342
Db	111	ttptcaattpttttaapttaatttaavtaatsaettatat-atstpttttpttttt	169
Qy	343	AEPGCTTSQVPTTEVISTAPVQMPFAESTGCTPEKVPVSEVNGTTLAEMSTPEATGTP	402
Db	170	atttvtpttaatttdtttaatttaatttaatttaatttaatttaatttaatt	229
Qy	403	AEVSVIWSLSTGTAQVTTTEWEVETARELPPEPEGPDASSIMSTESITGSLGLDCTA	462
Db	230	t	230

Aug 27 08:00

US-08-231-565A-27.rpt

15

Oy 463 T 463

Search completed: Tue Aug 27 08:12:37 1996
Job time : 102 secs.

WVWPSRELL (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 08:09:36 1996; MaaPar time 16.82 Seconds

Tabular output not generated.
683.05% Million cell updates/sec

Title: >US-08-231-565A-27

Description: (1-661) from US08231565A.pep

Perfect Score: 4698

Sequence: 1 MDVLKRLGLHLAVIGALLA.....RIFCSPICGENSPILLGQGV 661

Scoring table: PAM 150

Gap 11

Searched: 49340 seqs, 17385503 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9

Statistics: Mean 52.781; Variance 121.768; scale 0.433

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	4586	97.6	668	6	PM17_HUMAN	PMEL 17 PROTEIN PRECU	0.00e+00
2	206	4.4	400	5	MU1_XENLA	INTEGUMENTARY MUCIN A	6.74e-13
3	205	4.4	865	2	CPN_DROME	CALPHOTIN.	9.36e-13
4	169	3.6	797	8	VGLX_HSVB	GLYCOPROTEIN X PRECUR	8.77e-08
5	169	3.6	907	8	VGP3_EBV	ENVELOPE GLYCOPROTEIN	8.77e-08
6	167	3.6	1367	1	AMVH_YEAST	GLUCOAMYLASE SL/S2 PR	1.62e-07
7	164	3.5	600	7	SP96_DICDI	SPORE COAT PROTEIN SP	4.02e-07
8	163	3.5	2274	5	MUC2_HUMAN	MUCIN 2 (INTESTINAL M	5.44e-07
9	156	3.3	662	5	MU1_XENLA	INTEGUMENTARY MUCIN C	4.41e-06
10	148	3.2	670	8	VG50_HSVI1	HYPOTHETICAL GENE 50	4.58e-05
11	148	3.2	671	1	ALYS_ENTFA	AUTOLYSIN (EC 3.5.1.2	4.58e-05
12	149	3.2	739	6	OCT1_CHICK	OCTAMER-BINDING TRANS	3.43e-05
13	149	3.2	743	6	OCT1_HUMAN	OCTAMER-BINDING TRANS	3.43e-05

14	144	3.1	474	9	VTP3_TTVIV	VIRAL PROTEIN TPX.	1.44e-04
15	147	3.1	742	1	BAL_HUMAN	BILE-SALT-ACTIVATED L	6.11e-05
16	145	3.1	1251	9	Y0U3_CAEEL	HYPOTHETICAL 133.5 KD	1.09e-04
17	141	3.0	224	9	Y091_NPVAC	HYPOTHETICAL 24.1 KD	3.38e-04
18	141	3.0	542	6	PSGL_ONCMY	APOLYSIALOGLYCOPROT	3.38e-04
19	142	3.0	610	5	MUC4_HUMAN	MUCIN 4 (TRACHEOBRONC	2.55e-04
20	141	3.0	796	9	YSSA_CAEEL	HYPOTHETICAL 84.3 KD	3.38e-04
21	140	3.0	1025	7	SLAP_CAUCR	S-LAYER PROTEIN (PARA	4.48e-04
22	138	2.9	398	5	MU1_XENLA	INTEGUMENTARY MUCIN B	7.84e-04
23	137	2.9	567	2	CH13_CANAL	CHITINASE 3 PRECURSOR	1.04e-03
24	136	2.9	605	9	YHC8_YEAST	HYPOTHETICAL 63.8 KD	1.37e-03
25	137	2.9	699	8	VGLG_HSV2	GLYCOPROTEIN G.	1.04e-03
26	134	2.9	705	4	GUN6_DICDI	ENDOGLUCANASE PRECURS	2.37e-03
27	138	2.9	725	1	AGAL_YEAST	A-AGGLUTININ ATTACHEME	7.84e-04
28	132	2.8	173	4	HA34_BRELK	HAM34 PROTEIN.	4.09e-03
29	132	2.8	259	5	MSP8_EIMAC	MEROZOITE SURFACE PRO	4.09e-03
30	132	2.8	353	7	SUC3_RAT	SYNDECAN-3 (N-SYNDECA	4.09e-03
31	130	2.8	361	6	PIR3_YEAST	PIR3 PROTEIN PRECURSOR	7.04e-03
32	130	2.8	573	1	C114_MOUSE	CELL SURFACE ANTIGEN	7.04e-03
33	131	2.8	630	5	MUC1_MOUSE	MUCIN 1 PRECURSOR (PO	5.37e-03
34	128	2.7	400	5	LEUK_HUMAN	LEUKOSIALIN PRECURSOR	1.20e-02
35	129	2.7	676	4	ICP0_HSVBJ	TRANS-ACTING TRANSCRI	9.21e-03
36	126	2.7	1609	9	YCA9_YEAST	HYPOTHETICAL 166.0 KD	2.05e-02
37	129	2.7	1849	4	IGA4_HAETM	IMMUNOGLOBULIN A1 PRO	9.21e-03
38	127	2.7	1854	2	CIPA_CLOTM	CELLULOSOMAL SCAFFOLD	1.57e-02
39	127	2.7	2441	2	CBP_MOUSE	CREB-BINDING PROTEIN.	1.57e-02
40	123	2.6	338	7	SED1_YEAST	SEDI PROTEIN PRECURSO	4.51e-02
41	123	2.6	439	9	XP2_XENLA	SKIN SECRETORY PROTEI	4.51e-02
42	124	2.6	713	8	TUP1_YEAST	GLUCOSE REPRESSION RE	3.47e-02
43	124	2.6	767	1	AMVH_SACDI	GLUCOAMYLASE S1 PRECU	3.47e-02
44	123	2.6	1150	1	APMU_PIG	APOMUCIN (MUCIN CORE	4.51e-02
45	123	2.6	1162	8	TCNA_TRYCR	SIALIDASE (EC 3.2.1.1	4.51e-02

ALIGNMENTS

RESULT	1	HUMAN	STANDARD;	PRT;	668 AA.
ID	PM17_HUMAN				
AC	P40967;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)				
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92021023.				
RA	KWON B.S., CHINTAMANENI C., KOZAK C.A., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,				
RA	KIM K.-K.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).				
CC	-/- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	-/- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	-/- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; M77348.				
DR	MIM; 155550; 11TH EDITION.				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23	POTENTIAL.	
FT	CHAIN	24	668	PMEL 17 PROTEIN.	
FT	TRANSMEM	575	595	POTENTIAL.	
FT	TRANSMEM	603	623	POTENTIAL.	

	FT	DOMAIN	315	392	3 X 26 AA	TANDEM REPEATS.
	FT REPEAT	315	340	1.		
	FT REPEAT	341	366	2.		
	FT REPEAT	367	392	3.		
	FT CARBOHYD	81	81	POTENTIAL.		
	FT CARBOHYD	106	106	POTENTIAL.		
	FT CARBOHYD	111	111	POTENTIAL.		
	FT CARBOHYD	321	321	POTENTIAL.		
	FT CARBOHYD	568	568	POTENTIAL.		
	SO SEQUENCE	668 AA:	70992 MW:	2510968 CN:		

Query Match 97.6%; Score 4586; DB 6; Length 668;
Best Local Similarity 98.5%; Pred. No. 0.00e+00;
Matches 658; Conservative 0; Mismatches 3; Indels 7; Gaps 1;

Db	1	mdlvklrcllh	lavigallav	gatkyp	rnqdwlgv	srglrkt	kawnrqlp	ewtca	qrl	ldc	60						
Qy	1	MDLVKRCLLH	LAVIGALLAV	GATKYP	RNQDWLGVS	RGLRKT	KAWNRLQ	P	EWTEA	QRLDLC	60						
Db	61	wrggvs	lkvsndqpt	liganas	fsialnf	pgsqkv	lpdqgv	ivwnnt	iings	qvwgqgp	120						
Qy	61	WRGGVS	LKVSNDQPT	LIGANAS	FSIALNF	PGSQKVL	PDQGVIV	WNNTI	INGSQ	VWGGQP	120						
Db	121	vypqet	ddaci	fpdggp	pcpsqsw	qkrsf	vyvwk	twgqv	wlgvp	svglis	igt	ramlg	180				
Qy	121	VYPQET	DDACIF	PDGGPC	PCPSQSR	FSVYWK	TWQVW	QVWLG	VPVSG	LISIGT	GRAMLG	180					
Db	181	thtmevt	vyhrrgs	vyplah	ssafft	itdq	pfs	vsvsq	lral	dgnkh	flrn	qpltf	240				
Qy	181	THTMEVT	VYHRRGSR	VYPLAH	SSAFTIT	TDQVPF	SVVSQ	LRALD	GNKHF	L	RNQPLTF	240					
Db	241	alqlhd	pegylae	adlsy	twdfg	dsagtl	israp	wthyle	pgp	vtacv	qva	ipits	300				
Qy	241	ALQLHDP	SGYLAEAD	LSYTWDF	CDSSG	TLISRA	LWTHY	LEPCF	VTAQ	VVLQAA	IPLTS	300					
Db	301	cqgs	svp	qgtcd	ghrpt	aeapnt	agq	vptev	vglt	pgq	apta	esp	qst	sv	qpt	tevis	360
Qy	301	CQSSVP	QGTCDGHR	P	AEAPNT	AGQVPT	TEVGT	PGQAP	TAEP	SC	TT	SVQPT	TEVIS	360			

361	Db	tapvcmptaaestgmtpekvpsvsmgdtllaemstpeatgmtpaevsvlwgtdtaaqvtt	420
361	Qy		420
361	Qy		420
361	Qy		420
421	Db	tewvettarelpipepegdassimsteatslqplldgtatlrllvkrvpvlcdvlyry	480
421	Qy		480
421	Qy		480
481	Db	gafsvtldivqgieaeailqavpsgegdafeiltvscqglpkpeacmeispsacppaql	540
481	Qy		540
481	Qy		540
541	Db	cpvlpapacqvlwhqlkqgstytcnlvsadtnslavvstqlimpvpgilltqaeagl	600
541	Qy		593
601	Db	qgvrlivgillvmavvlasliiyrllrmkqdfsvpqlphsseshwlrprifescpi	660
601	Qy		660
601	Qy		660
654	Qy	qgvplivgillvmavvlasliiyrllrmkqdfsvpqlphsseshwlrprifescpi	653
661	Db		668
661	Qy		661
661	Qy		661

[illegible]

US-08-231-565A-27 JSP

US-08-231-565A-27.rsp

421	fkapeattstptIntgtfadpntttgtpsthyvntnltpaatgctgstadvstadvstptpag	480
262	FGDSGGLTSLRALVTHTYLEPCVTA-QVVLQAAIPLTSCGS-PVFGTTDGHRRPTAEA	319
481	ttesgasvptpspswdngteskapdm-tsstsvtttptnatsptpavttptnatsptp	539
320	PNTTAGVQ-PTTEV-VGTFPGQAPTAEPSTGTSVQVPTTEVISTAP-VQMTAESVGMTP	376
540	avttptnatsptlgktsptsavttptnatsptlgktsptsavttptnatsptlgkts	599
377	E-KVPSEVMGTTLAEMSTPEATGMTPAEVSIVLUGTTAAQVTTTPEWVETTELPIPE	435
600	ptsavttptnatsptgvtgtsqcanathtl	530
436	PEGPDASSI-MSTESITGSLGLLDGTA-TL	464

RESULT	6	
ID	AMVH YEAST	STANDARD; PRT; 1367 AA.
AC	P08640; P08068;	
DDT	01-AUG-1988 (REL. 08, CREATED)	
DDT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)	
DDT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)	
DE	GLUCOAMYLASE SI/52 PRECURSOR (EC 3.2.1.3) (GLUCAN 1,4-ALPHA-	
DE	GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE).	
GN	STAI OR STA2 OR WAL5 OR YIR019C.	
OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).	
OC	EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RC	STRAIN=S288C / AB9712;	
RR	BARELL B.G., BADCOCK K., BANKIER A.T., BOWMAN S., BROWN D.,	
RR	GURCHER C.N., CONNOR R., COPEY T., DEAR S., DEVLIN K., FRASER A.,	
RR	GENTILES S., HAMLIN N., HORSNELL T.S., HUNT S., JAGELS K., JONES M.,	
RR	LOUIS E., LYE G., MOULE S., MOULE T., ODELL C., PEARSON D.,	
RR	RAJANDREAM M.A., RILES L., ROWLEY N., SKELTON J., SMITH V.,	
RR	WALSH S.V., WHITEHEAD S.;	
IRL	SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.	
RN	[2]	
RN	SEQUENCE OF 1-242 AND 762-1331 FROM N.A.	
RR	MEDLINE; 87194600.	
RR	YAMASHITA I., NAKAMURA M., FUKUI S.;	
RR	J. BACTERIOL. 169:2142-2149(1987).	
RN	[3]	
RN	SEQUENCE OF 1-31 FROM N.A.	
RC	STRAIN=SPX101-1C;	
RC	MEDLINE; 89031230.	
RR	PARDO J.M., IANEZ E., ZALACAIN M., CLAROS M.G., JIMENEZ A.;	
RR	FEBS LETT. 239:179-184(1988).	
CCC	-/- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL 1,4-LINKED ALPHA-D-	
CCC	GLUCOSE RESIDUES SUCCESSIVELY FROM NON-REDUCING ENDS OF THE CHAINS	
CCC	WITH RELEASE OF BETA-D-GLUCOSE.	
CCC	-/- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES.	
DR	EMBL; Z47047; Z47047.	
DR	EMBL; M16164; M16164.	
DR	EMBL; M16165; M16165.	
DR	EMBL; X13857; X13857.	
DR	PIR; B26877; B26877.	
DR	PIR; A26877; A26877.	
DR	PIR; S48478; S48478.	
DR	LISTA; SC01218; STAI.	
DR	SGD; L0002104; STAI.	
DR	PROSITE; PS00820; GLUCOAMYLASE.	

Query Match 3.6%; Score 169; DB 8; Length 907;
Best Local Similarity 23.2%; Pred. No. 8.7/e-08;
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;

KW HYDROLASE; GLYCOSIDASE; POLYSACCHARIDE DEGRADATION; GLYCOPROTEIN;
KW SIGNAL; MULTIGENE FAMILY.

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 1367 GLUCOAMYLASE S1/S2.
FT DOMAIN 210 1367 SER/THR-RICH.
FT CARBOHYD 817 817 POTENTIAL.
FT CARBOHYD 874 874 POTENTIAL.
SQ SEQUENCE 1367 AA; 136110 MW; 12881462 CN;

Query Match 3.6%; Score 167; DB 1; Length 1367;

Best Local Similarity 24.8%; Pred. No. 1.62e-07;

Matches 36; Conservative 39; Mismatches 69; Indels 1; Gaps 1;

Db 574 Lessapvptssstessapvptssstessapvptssstessapvptssstessapvptssstess 633

Qy 299 TSCSSPVPCTDGRPTAEAPNTAGQVPTTEVVGTPCQAPTAESGTSVQVPTTEV 358

Db 634 -esapvptssstessapvptssstessapvptssstessapvptssstessapvptssstess 692

Qy 359 ISTAPVQMPETAESTGMPKPVSEVMGTTLAEMSTPEATGMPAEVSVVLSGTTAAQV 418

Db 693 apvtsstessapvptssstess 717

Qy 419 TTTEWVETARELP IPEPEGPDASS 443

RESULT 7

ID SP96 DICDI STANDARD; PRT; 600 AA.
AC P14328;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE SPORE COAT PROTEIN SP96.
CN COTA.

OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).

OC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;

OC EUKARYOTA; DICTYOSTELIA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RX MEDLINE; 90067962.

RA FOSNAUGH K., LOOMIS W.F.;

RL NUCLEIC ACIDS RES. 17:9489-9489(1989).

DR EMBL; X16491; X16491.

DR PIR; S07638; S07638.

DR DICTYDB; DD03007; COTA.

KW SPORULATION; REPEAT.

FT REPEAT 189 197

FT REPEAT 225 233

FT REPEAT 302 310

FT REPEAT 399 407

SQ SEQUENCE 600 AA; 59589 MW; 2017994 CN;

Query Match 3.5%; Score 164; DB 7; Length 600;

Best Local Similarity 18.6%; Pred. No. 4.02e-07;

Matches 32; Conservative 64; Mismatches 74; Indels 2; Gaps 2;

Db 424 gsdssalqstssssgssaa 483

Qy 284 GCVTAQVWVLAAPLITSCSSSPVCTDGRPTAEAPNTAGQV-PTEVVGTPCQ-AP 341

Db 484 ssspsss 543

Qy 342 TAEPSTGTSVQVPTTEVISTAPVQMPETAESTGMPKPVSEVMGTTLAEMSTPEATGMT 401

Db 544 ttaatt 595

Qy 402 PAEVSIVVLSGTTAQVTTTEWVETARELP IPEPEGPDASSINSTESITGS 453

RESULT 8

ID MUC2 HUMAN STANDARD; PRT; 2274 AA.

AC Q02617;

DT 01-JUN-1994 (REL. 29, CREATED)

DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)

DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)

DE MUCIN 2 (INTESTINAL MUCIN 2) (FRAGMENTS).

GN MUC2 OR SMOG.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE OF 1-1270 AND 1291-2274 FROM N.A.

RC TISSUE-COLON;

RX MEDLINE; 93016075.

RA GUM J.R. JR., HICKS J.W., TORIBARA N.W., ROTHE E.-M., LAGACE R.E.,

RA KIM Y.S.;

RL J. BIOL. CHEM. 267:21375-21383(1992).

RN [2]

RP SEQUENCE OF 718-1290 FROM N.A.

RX MEDLINE; 91358717.

RA TORIBARA N.W., GUM J.R. JR., CULHANE P.J., LAGACE R.E., HICKS J.W.,

RA PETERSEN G.M., KIM Y.S.;

RL J. CLIN. INVEST. 88:1005-1013(1991).

CC -!- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND

CC OTHER MUCUS MEMBRANE-CONTAINING ORGANS.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR

CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).

CC -!- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,

CC BRONCHUS, CERVIX AND GALL BLADDER.

CC -!- SUBUNIT: MULTIMERIC.

CC -!- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND

CC VARIES AMONG DIFFERENT ALLELES.

DR EMBL; M74027; M74027.

DR EMBL; M94131; M94131.

DR EMBL; M94132; M94132.

DR HSSP; P15445; 1PSH.

DR MIM; 158370; 11TH EDITION.

DR PROSITE; PS00022; EGF.

KW GLYCOPROTEIN; REPEAT; POLYMORPHISM.

FT NON_TER 1 1

FT DOMAIN 776 1122 APPROXIMATE REPEATS.

FT REPEAT 776 791 1.

FT REPEAT 792 807 2.

FT REPEAT 808 823 3.

FT REPEAT 824 839 4.

FT REPEAT 840 846 5.

FT REPEAT 847 853 6.

FT REPEAT 854 869 7A.

FT REPEAT 870 892 7B.

FT REPEAT 893 908 8A.

FT REPEAT 909 931 8B.

FT REPEAT 932 947 9A.

FT REPEAT 948 971 9B.

FT REPEAT 972 987 10A.

FT REPEAT 988 1010 10B.

FT REPEAT 1011 1026 11A.

RESULT	11	
ID	ALYS_ENTFA	STANDARD;
		PRT; 671 AA.

RESULT	12
ID	OCT1 CHICK STANDARD; PR7; 739 AA.
AC	P15143;
DT	01-APR-1990 (REL. 14, CREATED)
DT	01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT	01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE	OCTAMER-BINDING TRANSCRIPTION FACTOR 1 (OTF-1) (NF-A1).
GN	OTF1 OR OCT1.
OS	GALLUS GALLUS (CHICKEN).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE; GALLIFORMES.
CC	(1)
RN	SEQUENCE FROM N.A.
RP	MEDLINE; 90138945.
RX	RA PETRYNIAK B., POSTEMA C.E., MCCORMACK W.T., RA THOMPSON C.B.;

FT REPEAT 724 734
SQ SEQUENCE 742 AA; 78345 MW; 3107924 CN;
Query Match 3.1%; Score 147; DB 1; Length 742;
Best Local Similarity 24.6%; Pred. No. 6.11e-05;
Matches 42; Conservative 50; Mismatches 67; Indels 12; Gaps 12;
Db 567 eatpvptgdsdapvptgdsagppvptgdsagppvptgdsagppv 626
Qy 282 EPGVTHAQVWLQAA-IPLTS-CGSSFPVGTDD-GHRPTAEAPNTTACQVPTTEVVGTTTG 338
Db 627 -pptgd-sagppvp-ptgds-gappvp-ptgds-g-appvpptgdagppvpvptgdsagap 679
Qy 339 QAPTAEPSTTSVQVPTTEVISTAPVQMPVTAESTGTMPEKVPVSEVMGTTLAEMSTPEAT 398
Db 680 pvvptgdsagppvptgdsdapvptgdsag-appvpptgdsagppvptgd 729
Qy 399 GMTPA-EVSIVVLGTTAAQVTTTEWVETTARELPPEPEGPDASSIMSTE 448

Search completed: Tue Aug 27 08:10:37 1996
Job time : 61 secs.

Db	1342	gccagctcaatcgtctcaggaagtattacaggttccctggcccccctgctggatggt	1401
Qy	1321	GCNSNWSNATHATGNSNACNARKSNATHACNGNSNYTNGCNCNYTNYNGAYGN	1380
Db	1402	acagccacctaagctggtgaagagacaagtcctcctggattggtctctgcatcatat	1461
Qy	1381	ACGNCNACNYTNGNYTNGTNAARMGNCARGTNCNCNYTNGAYTGTNYTAYMGNTAY	1440
Db	1462	ggtctcttttcgcaccctggacatgctccaggtattgaaagtgcagagatcctgcag	1521
Qy	1441	GCNSNTTWSNCTNACNYTNGAYATGTCNARGNATGARGNSNCNARATHYTNCAR	1500
Db	1522	gctgtgccctccggtgaggggagtcatttgagctgactgtctcctgccaggcggtg	1581
Qy	1501	CGCTGNCNWSNCGNARGANGCAYGNTTYGARYTNACNCTGNSNTGYCARGGNGNYN	1560
Db	1582	cccaaggaagcctgcatgagatctcctccaggtgcccagcccctgccagcgctg	1641
Qy	1561	CONAARGCNGTGYATGARGATHHSNWSNCGNCTGYCARGCNCNGCNCARGNYN	1620
Db	1642	tgccagctgtgctaccagccagcctgccagctggttctgcacagatactgaaggt	1701
Qy	1621	TCYCARCNGTNYTNCNWSNCCNGTGYCARYTNGTNYTNCAYCARATHYTNARGCN	1680
Db	1702	ggctcggggagacatacctcctcaatgctctcctgctgataccacagcctggcagtgctc	1761
Qy	1681	GCNSNWSNACNTATYTYTNAAYTWSNCTNYTNGCNGAYACNAAYWSNYTNGCNGTGN	1740
Db	1762	agcaccagcttcatcgtcgtgaagaagcagcctgtggcaggttcctgctgacgtg	1821
Qy	1741	WSNACNARYTHATATGCCNCGNARGCNGGNYTNGCNGCARGTNCNCTNYTHGTN	1800
Db	1822	ggcatctgctggtgagctgctgctcctgctcctcctgctcctgataatagccagact	1881
Qy	1801	GGNATHYTYTNGTNYTATGGCNGTNGTNYTNGCNSYTNATHATYTMNGNMGNYN	1860
Db	1882	atgaagaagacttccgtaccagctgccaatagcagagcagcactgctgctgctc	1941
Qy	1861	ATGAARCAGAYTYSNCTGNCNARYTNCNCAYWSNWSNWSNCAYTGTGTYNMGNYN	1920
Db	1942	ccccgatctcctgctgctcctcctgctgagagatagccccctcctcagtgggcagcag	2001
Qy	1921	CONMGNATHYTYGWSNTGYCNCNATHCGNARGAAYWSNCCNYTNYTWSNNGNCAR	1980
Db	2002	gt 2003	
Qy	1981	GT 1982	
RESULT	2		
LOCUS	HUMGPMS	2114 bp	mRNA
DEFINITION	Human 95 kD melanocyte-specific secreted glycoprotein mRNA, 3' end.		
ACCESSION	M32295		
NID	g183559		
KEYWORDS	melanocyte-specific secreted glycoprotein.		
SOURCE	Human melanoma cell line, cDNA to mRNA, clone 8.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 2114)		
AUTHORS	Vogel, A.		
TITLE	Sequence of a melanocyte specific secreted glycoprotein		
JOURNAL	Unpublished (1990) St. Louis Univ., St. Louis, MO 63104		
COMMENT	Draft entry and computer-readable sequence for [1] kindly submitted		

by A.Vogel, 23-FEB-1990.			
NCBI gi:	183559	Location/Qualifiers	
source	1..2114	/organism="Homo sapiens"	
CDS	<1..2028	/note="melanocyte-specific secreted glycoprotein; NCBI gi: 386754"	
		/codon_start=1	
		/db_xref="PID:g386754"	
		/translation="RGIRKNTWDLVLRCLLHLAVIGALLAVGATKVPKRDQWLGVSR	
		QLRTKAHRNQLYFETWEAQRDLCWRGQVSLKVSNDPDLICNASFIALNPFCSQK	
		VLPDQVWVANNVTIINGSQVWGQVPYPOETDDACIFPDGCGPCSGWSQKRSFYVM	
		KTMGQVQVGLGPGVSGISCTGRAMLGHTMEVTVYHRRGSRSYVP LAHSSSAFTTD	
		QVPSVSQIRALDGNKHFIRNQLPTEALQLDHPSGLAEADLSYTMDFGDSSTGL	
		ISRALVYTHYLEPGVTAQVVIQAAIPLTSCGSSPVPDTHDRPFAEPNTAGQV	
		VPSEGDAFELTVSCQGLPKAECSISPPCQPPAQRCLQVLPSPAQDLVHQILK	
		GGSCYCLNVLADTMSLAVSTOLIMPQGEAGLGQVPLIVGLILVLMVLAHLIR	
		RRLMKQDFSVLPQLPHSSSHLRLPASALVPLVLRAPSSVSGSRSEYSHHML"	
BASE COUNT	469 a	586 c	575 g 484 t
ORIGIN			
	Query Match	96.2%; Score 7334; DB 55; Length 2114;	
	Best Local Similarity	57.0%; Pred. No. 0.00e+00;	
	Matches 1129; Conservative 416; Mismatches 436; Indels 1; Gaps 1;		
Db	22	atggatcgtgctcaaaagatgccttccttcatttgctgctgatatgctgcttgctgctgct	81
Qy	1	ATCGAYTYTNGTNYTNAARMGNTGYTYTNCAYTYTNGCNGTNAHCGNCGNYTNYTNGCN	60
Db	82	gtgggggctacaaaagaccagaaacagcagctgctgctgtctcagaagcaactcaga	141
Qy	61	GTGNGCNGACNAARGTNCNMGNAAYCARGAYTGGYTNGCNGTWSNMGNCARYTNMGN	120
Db	142	accaaagcctggaacagcagcgtgtatccagagtggacagaagccagcagacttgctgc	201
Qy	121	ACNAAARGCNTGGAYTMGNARYTNTAYCNGARTGGACNGARCNARMGNYTNGAYTGY	180
Db	202	tgagagagtggtcaagtgctccctcaaggtcagtaatgatggcctacactgattggtgca	261
Qy	181	TGGMGNGCNGCARGTNSNYTNAARGTNSNAAAYGAYGNCNACNYTNATHGNGCN	240
Db	262	aatgctcctctctctgcttggaacttcctcctggaagccaaaagtgattgccagatggg	321
Qy	241	AAYGNSNTTWSNATHGNYTNAAYTYCCNCGNWSNCARAARGTNYTNCNGAYGN	300
Db	322	caggttatcgtgggtcaacaataccatcatcaatggagccagcagtggtgggagagcagcca	381
Qy	301	CARGTNATHTGGGTNAAYAYACNATHATHAAYGNGNSNARGTNTGGGNGNCARCN	360
Db	382	gtgtatccccaggaactgacgactgctgctcctcctcctgctgctggaccttgcccatct	441
Qy	361	GTNTAYCNCARGACNARGAYGCNTGYATHTTYCNGAYGNGGNCNCNYTCNWSN	420
Db	442	ggctccttggtctcagaagagagcttggtttatgctggaagacctggggccaactcag	501
Qy	421	GCNWSNTGNSNCARARMGNSNTTYCTNTAYCTNTGGAARACNTGGGNCARTAYTGG	480
Db	502	caagtctcagggggcccgctgctgggctgagcattggagacaggcaggcaatgctgggc	561


```
RESULT 6
ID HSMEL17S2 standard; DNA; PRI; 2817 BP.
AC U20093;
DT 01-JAN-1996 (Rel. 46, Created)
DT 01-JAN-1996 (Rel. 46, Last updated, Version 1)
DE Human melanocyte-specific (pmel 17) gene, exons 2-5, and complete
DE cds.
KW Homo sapiens (human)
OS Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-2817
RA Kwon B.S., Kim K., Heng H.H., Shi X.M., Tsui L., Lee Z.H.,
RA Youn B., Pickard R.T.;
RT "Genomic organization and FISH mapping of human pmel 17, the
RT putative silver locus";
RL Unpublished.
RN [2]
RP 1-2817
RA Kwon B.S.;
RT ;
RL Submitted (05-JAN-1995) to the EMBL/GenBank/DBJ databases.
RL Byoung S. Kwon, Indiana University School of Medicine, Microbiology
RL and Immunology, 635 Barnhill Drive, Indianapolis, IN 46202, USA
CC NCBI gi: 1142634
FH Key Location/Qualifiers
FH 1..2817
FH FT source
FH FT /clone_lib="lambda FIX II from Strategene"
FH FT /organism="Homo sapiens"
FH FT /chromosome="12"
FH FT /map="12pter-q21"
FH FT join(019491:1111...1186,67..459,853..1014,1266..2396,2499..
FH FT 2722)
FH FT /gene="pmel 17"
FH FT /codon_start=1
FH FT /product="Pmel 17"
FH FT /db_xref="PID:g1147636"
FH FT /translation="MDLVKRLHLAVIGALLAVGATKVPKRNQDWLGVSRLRTKAM
N RQLYPENWEAQRLDCHRGQVSLKVSNDGPTLLIGANAFSIALNPPCSQKVLPGQVI
W VNNTIINGSQVGGQVPVQETDDACIFPDGQPCPSGWSQKRSFVYVWKTWQVYQV
L GGPVSGLSIGTGRLMLGTHMEVTVYHRRGSYSVPLAHSSSAFTIIDQVPFSVSQ
L RALDGNKHFLRNQPLTFALQLDHPSCYLAELLSYTWDFGDSGLTISRAPVWTHY
L EPGVTAQVWLQAAIPLTSCGSSVPVGTTHDHRPTAEAPNTTAQVPTTEWVGTTTPGQ
L PTAEPSTTSVQVPTTEVLSIAPVQMPTEASTGHTPEKVPVSEVWGTTLAEMSTPEAT
A MTPAEVSIWLSGTTAAQVTTTEWVETARELP IPEPEGFADASSIMSTESITGSLGL
G DGTATURLKQVPLDCVLYRYGSFVTLDIQVIESAEILQAVPSCGDAPFLVSC
L GGLPKEACWEISSFGQPPAQRQLQVPLSPACQLVLHLQLKGGSGTYCLNLSLADTN
Q LAVVSTQLIMPQGEAGLGQVPLIVGILLVMAVVLASLIYRRRLMKQDFSVPLPHSS
S HWLRALPRIFCSCP IGENSPILSGQV"
```

```
FT exon 67..459
FT /gene="pmel 17"
FT /number=2
FT exon 853..1014
FT /gene="pmel 17"
FT /number=3
FT exon 1266..2396
FT /gene="pmel 17"
FT /number=4
FT exon 2499..>2722
FT /gene="pmel 17"
FT /number=5
SQ Sequence 2817 BP; 627 A; 772 C; 739 G; 679 T; 0 other;

Query Match 54.7%; Score 4170; DB 10; Length 2817;
Best Local Similarity 57.1%; Pred. No. 0.00e+00;
Matches 652; Conservative 224; Mismatches 265; Indels 0; Gaps 0;

Db 1257 caatccagaccagtgctttctccgtgagcgtgccagttgcctccagctccatgaccaca 1316
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 623 CNATHACNGAYCARGTNCNTTYSNGTNSGNTWNSNCARYTNMGCONGTNGAYGNG 682

Db 1317 ggaacaagcacttccagaaatcagcctcagccttgcctccagctccatgaccaca 1376
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 683 GNAAYARCAATTTTNNGMAAYCARCONYTACNTTTCNTYTCARNTNCAYGAYCONW 742

Db 1377 gtgctatctggctgaagtgacctctcctacacctgggaccttggagacagtgaggaa 1436
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 743 SNGCNTAYTNGCNGCNGCNGAYTNSNTAYACNTGGGAYTTTGGCAYTWSNWSNGNA 802

Db 1437 cctgatctctcgccacactgtggtcactcactacactgagcctgcccacgtcactg 1496
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 803 CNYTNATHWSNMGCONGTNGTNGTACNCACTATYTTNGARCONGCONGTNAONG 862

Db 1497 cccaggtggctcagagtgccattcctcctcctcctcctcctcctcctcctcctcag 1556
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 863 CNGCARTGTNTTCARCGNGCNGTTCNTYACNTGACNTGACNTGACNTGACNTGACNTG 922

Db 1557 gcacacagatgggcacagggcacaactgcagagggccctaacacacacagctggccagtgc 1616
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 923 GNACNACNGAYGNCAYMGNCNACNCGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGTNC 982

Db 1617 ctactacagaagtgtgggtactacacctgtcagggcccaactgcagagccctctggaa 1676
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 983 GNACNACNGARTGTNGNACNACNCGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGNA 1042

Db 1677 ccacatctgtcaggtgcacacacactgaagtcataagcactgcacctgtgcagatgccaa 1736
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 1043 CNACNWSNNGTNCARTGCCNACNACNARGTNCNACNACNCGCNGCNGCNGCNGCNGCNGNA 1102

Db 1737 ctgcagagacacaggtatgacacctgagaagtgccagtttccaggtcatgggtacca 1796
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 1103 CNGCNGARMSNACNGNATGACNCGCARAARTTCCNGTNSNGARTNATGGGNACNA 1162

Db 1797 cactggcagagatgcaactccagaggtacaggtatgacacctgcagaggtatacaattg 1856
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 1163 CNYTNGCNGARATGMSNACNCGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGTNSNATHG 1222

Db 1857 tgggtctttctggaaccacagctgcacaggtcaacaactacagagtggtggtagaccacag 1916
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 1223 TNGTNTWSNWSNGNACNACNCGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNGCNG 1282

Db 1917 ctadagagctacctatccctgagcctgaaggtccagatgccagctcaatcatgtctacgg 1976
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
```


[illegible]

JOURNAL	preferentially expressed in retinal pigment epithelium					
COMMENT	Curr. Eye Res. 55, 657-662 (1992)					
FEATURES	NCBI gi: 162770 Location/Qualifiers					
source	1..1584					
	/organism="Bos taurus"					
	/dev_stage="calf"					
	/sequenced_mol="cDNA to mRNA"					
	/tissue_type="retinal pigment epithelium"					
	/tissue_lib="Unizap XR"					
CDS	1..1477					
	/partial					
	/gene="RPE1"					
	/note="ORF; NCBI gi: 162771"					
	/codon_start=2					
	/product="retinal pigment"					
	/db_xref="pid:g162771"					
	/translation="QYMQVLGPGVSGLSIGTDKAMLGTYNMVEVTVVHRRSGSYPLA HSSAFITIDQVPFSVSVSQIQADGNRKRLRKQPIPTALQLHDPSGYLAGADLSYT WFGDSTGLISRALTHTHYTHIESGPVTAQQVAAPLPLSCGSPPVGTTDRHVHTTA EAPCTTAGQVPTVMCGTTPGGVTAEPQTGVPTTEPDGTTPAQVATSKVLST PVMPTAKATGRTEXSTTESGTTVTQGTTELVEVTTAGETVSETPEPAGNTSSFMT EGTAGSLSPTRDTATVILXEQRAPLDCLVRYGSESFLDIVSIEASELQAQVSSSE GDPALEYSQCGGLPKEDACISSPCQAPQRICQPVPSPQACIQLDHQVLKGSGT YCLMVISLAQNS LAWSYTLVMPQGEARLRQAPLFVGIILLVITALLIASLIYYRRRLMK QSSEVPLPOLPHIGITQWLRLPWERSOPICEKPLLSGQQV"					
BASE COUNT	340 a 473 c 418 g 353 t					
ORIGIN						
	Query Match 38.4%; Score 2926; DB 33; Length 1584;					
	Best Local Similarity 49.2%; Pred. No. 0.00e+00;					
	Matches 746; Conservative 292; Mismatches 428; Indels 51; Gaps 23;					
Db	2	caatactggcaagtgtcgggggccagdtgctggactgagcatcggcacagacaaggca	61			
		: : :	: : :	: : :	: : :	: : :
Qy	472	CARTAYTGCCARTYTYNGNGNCGCNWTMSNGNYTNNSNAITHGNCNCGNMGNCN	531			
Db	62	atgctgggcacataaacatgaagtgaactgtaccacgccgggggtccacagatcat	121			
		: : :	: : :	: : :	: : :	: : :
Qy	532	ATGTYNNGNCAACACATGCAGTACNGTNTAYCAYMNGNMGNNWSMNWSNTAY	591			
Db	122	gtgccctcgctcaactccaagtccaacctcacattactgaccagtgccccttcctgtg	181			
		::: : :	::: : :	::: : :	::: : :	::: : :
Qy	592	GTNCCNTYNGNCAYSNWSNWSNGCNTTACNATHAGCAYCARGTCNCCNTTYNSGNT	651			
Db	182	agtgtgtcagctgcagccttgatggagaagaaacagccttcctgagaagcagcct	241			
		::: : :	::: : :	::: : :	::: : :	::: : :
Qy	652	WSNGTNSNCARYTNMGNGCYNTGAYGCGNGNAYAAACAAYTTTTYTMNGNAYCARCN	711			
Db	242	ctgacctttgccctccagctccaatgatccccagtgcttatgttgctgggctgacctttcc	301			
		::: : :	::: : :	::: : :	::: : :	::: : :
Qy	712	YTNACTYTGENTWCARYTNCAYGAYCMWSNGNTAYTYNCGNARGCNGAYTYNWSN	771			
Db	302	tacactggacttgttgacactacaggacccctgatctcgggcactcacgtgcaact	361			
		: : :	: : :	: : :	: : :	: : :
Qy	772	TAYACNTGGGAYTYGNGAYSNWSNWSNAGCNATYTNHWSNMNGNCTYNTGNTNGNACN	831			
Db	362	cacacttaactaagctctggcccaagtcaactgcacaggtggtgctcagagctgccatcct	421			
		: : :	: : :	: : :	: : :	: : :
Qy	832	CAYACNTAYTYNRCNCGNCGNACNACNARGCTNGTNTYTCARGCNGCNATHCNC	891			
Db	422	ctaactcctgtggctcctctccaagtccaaggcactacaagataggactgtgcacaactgca	481			
		::: : :	::: : :	::: : :	::: : :	::: : :
Qy	892	YTNACNWSNTGCGNWSNWSNCGNCTNCGNACNACNAGCAYGNCAYTMGNCNACNCGN	951			

[illegible]

RN	[1]
RP	1-673
RA	Bailin T., Lee S.T., Spritz R.A.;
RT	"Genomic organization and sequence of D12S53E (Pmel 17), the human
RL	homologue of the mouse silver (si) locus";
RR	Unpublished.
RN	[2]
RP	1-673
RA	Spritz R.A.;
RT	:
RL	Submitted (17-JUL-1995) to the EMBL/GenBank/DDBJ databases.
RR	Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of
WL	Wisconsin, Madison, WI 53706, USA
CC	NCBJ gi: 1125058
FH	Key
FH	Location/Qualifiers
FT	source
FT	1..673
FT	/organism="Homo sapiens"
FT	/chromosome="12"
FT	/map="12q13-q14"
FT	/note="DSEG number: D12S53E"
FT	exon
FT	128..241
FT	/number=2
FT	exon
FT	400..546
FT	/number=3
SQ	Sequence 673 BP; 186 A; 142 C; 179 G; 166 T; 0 other;
Query Match	7.5%; Score 570; DB 10; Length 673;
Best Local Similarity	56.0%; Pred. No. 9,74e-58;
Matches	89; Conservative 37; Mismatches 33; Indels 0; Gaps 0; O;
Db	387 ttcctccaggtggtcaagtgcctccaagtcagtaatgatggcctacactgattgg 446 : : : :
Qy	177 YTGTCGNGGGNCARGCTWNSYTNAAAGTWNWSNAAYGAYGNNCCACNTYNATHGG 236 : : : :
Db	447 tgcaaatgcctctcttattgccitgaacctcctcgaaagcaaagaatttgccaga 506 : : : :
Qy	237 NCGNAAVGCNMSNTYTWSNATHGCVYTNAAATTYYCNGCNWSCARAARGTNYTCNCA 296 : : : :
Db	507 tgggcaggttatctgggtccaacaataaccatacatg 545 : : : :
Qy	297 YGNCARGTNATHGGGTWTAAYAAYACGNATHAAYVG 335 : : : :
RESULT	15
LOCUS	HSPMEL05 1178 bp DNA PRI 19-DEC-1995
DEFINITION	Human melanocyte protein Pmel 17 gene, exon 11 and partial cds.
ACCESSION	U31799
NID	gl125061
KEYWORDS	.
SEGMENT	5 of 5
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 1178)
REFERENCE	Bailin,T., Lee,S.-T. and Spritz,R.A. Genomic organization and sequence of D12S53E (Pmel 17), the human homologue of the mouse silver (si) locus Unpublished (1995) 2 (bases 1 to 1178)
AUTHORS	Bailin,T.
TITLE	Direct Submission
JOURNAL	Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445

(TM)

Result	No.	Score	Query			ID	Description	Pred. No.
			Match	Length	DB			
	1	7374	96.7	2172.17	T02716	MART-1 melanoma anti	0.00e+00	
	2	7364	96.6	2115.16	Q96055	Sequence encoding imm	0.00e+00	
	3	6899	87.9	2131.17	T03760	Melanoma-specific imm	0.00e+00	
	4	177	2.3	1047.2	Q10572	Human Natriuretic Pep	3.52e-02	
	5	173	2.3	2351.2	Q10498	B-amyloidquefaciens s	6.56e-02	
	6	157	2.1	1065.17	Q94335	Degenerate Alteromona	7.46e-01	
	7	150	2.0	36.16	T05466	Sequence encoding imm	2.10e+00	
	8	147	1.9	755.1	Q03970	Clone P25b	3.24e+00	

DR P--PSDB; R84854.

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.

PS Disclosure; Fig 4A-4B; 184pp; English.

CC The nucleic acid encodes cDNA25, a melanoma antigen (MART-1)
CC which is recognized by T-lymphocytes. cDNA25 is a derivative of
CC the melanocyte-melanoma-specific antigen gp100 (see R84855).
CC Antigen cDNA25 is a source of immunogenic peptides (see R84199)
CC which are optionally modified (see R84200-R84211) and used in
CC medicaments, especially vaccines, for the treatment or prevention
CC (by immunization) of melanoma. Antibodies against cDNA25 and its
CC immunogenic peptides may be used in the detection and isolation
CC of the antigen from a sample, the detection of which is indicative
CC of a disease state (melanoma or metastatic melanoma).

SQ Sequence 2172 BP; 512 A; 594 C; 578 G; 488 T;

Query Match 96.7%; Score 7374; DB 17; Length 2172;
Best Local Similarity 57.0%; Pred. No. 0.00e+00;
Matches 1130; Conservative 416; Mismatches 436; Indels 0; Gaps 0;

Db 39 atggatctgggtctaaagaagctcttcattcgtgtgataggtgctttgctggct 98
Qy 1 ATGGAYTNGTNTYNAARMGNTGYTNTYNCAYTTCGCTNATHGGNGCNYTNTGCGN 60

Db 99 gtgggggtctaaagtaccgaagaaaccagagctggtgtgtctcaaggcaactcaga 158
Qy 61 CTNGGNGCNACNAARGTGTCNNGNAAACARGAYTGGYTGNGGTGNTWSMNGNCARYTNMGN 120

Db 159 accaaagcctggaacagcagctgtatccagagtggaagacagccagagcttgatgc 218
Qy 121 ACNAARCGTGGAAVMGNCARYTNTAYCCNGARTGGACNGARCCNARMGNTNGAYTGY 180

Db 219 tggagagtggtcaagtgccctcaagtcagtaatgatggctacactgattggtgca 278
Qy 181 TGGMNGNGNGCARGTNSNSYTNNAARGTNWSNAAAYGAYGCGCCNACNAYTNATHGGNGCN 240

Db 279 aatgcctctctctattgcttgaacttccttgaagccaaagattgcccagatggg 338
Qy 241 AAYCGNWSNTTYSNATHGCTNTAAATYTCGNGWSNCARAAAGCTNTNCNGAYCEN 300

Db 339 caggttatctgggtcaacaataccatcatcaatggagccaggtgtggggagacagcca 398
Qy 301 CARGTNATHGGGTNAAAYAAACNATHAAYGCGWSNCARGTWTGGGNGGNCARCCN 360

Db 399 gtgtatcccccaggaactgacgatgcctgeatcttccctgatggtggaccttgcccatct 458
Qy 361 GTNTAYCCNARGARACNAYGAYGCGTGYATHTYTCNAYGCGNGCNCCTGYCOWSN 420

Db 459 ggccttggtctcagaagaagctgtgttttatgtctggaagacctgggcccactactgg 518
Qy 421 GGNWSNTGCSNCARARMGWSNNTYGTNTAYGNTTGGAAACNCTGGCNCARTAYTGG 480

Db 519 caattctcagggggccagtgctcgggtgagcattgggacagcaggcgggcaatgctgggc 578
Qy 481 CARTTYTNGGNGCCNGCTNWSNGNYTNWSNATHGNCACNGMNGNCGNATGYTNGEN 540

Db 579 acaacacaccatggaagtactgtaccatccgcggggatcccgagagctatgtgctctt 638
Qy 541 ACNCAACNATCGARTNACNGTNTAYCAVMGNGMNGSNMNGWSNTAYGTNCNVTN 600

Db 639 gctcattccagctcagccttcaccattactgaccaggtgcttcttcctcgtgagcgtgccc 698
Qy 601 GCNCAYSNWSNWSGCGNTTYACNATHACNAYCARGTNCCTTYSNGTNGWSNGTWSN 660

Db 699 cagttgcccgttggtgatggagggaacaagcacttctctgagaaatcagcctctgaccttt 758
Qy 661 CARYTNNGCNYTNGAYGCGGNAAYAAARCAATYTYTNMGNAAYCARCCNYTNACTNTY 720

Db 759 ggcctccagctccatgaccccaagtgtctatctggctgaagctgaacctctctacacctgg 818
Qy 721 GCNYTNCARYTNCAYGAYCCNWSNGNTAYTYTNGCNGARGCNGAYTNTWSNTAYACNTGG 780

Db 819 gactttgagacagtagtggaaacctgatctctcgggcacttgtgtcactcactcatacttac 878
Qy 781 GAYTTYGNGAYGWSNWSNGNACNYTNATHWSMNGCGNTYTNCTNGTNGTNAACNAYACNTAY 840

Db 879 ctgagagctggcccagtcactgcccaagtggtctcgtcagagtgccattctcctcaccctcc 938
Qy 841 YTNARGCNGCNGCNGTNAACNGCARGTNGTNTNCARGCNGCNATHCCNYTNACNWSN 900

Db 939 tgtggtctctcccagttccaggccacacagatgggcacaggcccaactgcagaggccct 998
Qy 901 TGYGNSWSNWSNCCNGTNCNGNACNACNGAYGNCAYMGCNCCNACNGCARGCNGCNCN 960

Db 999 aaacacacagctggcccaagtgcctactacagaagttgtgggtactacacctggtcaggcg 1058
Qy 961 AAYACNACNGCNGCARGTNCNCCNACNACNGARGTNGTNGNACNACNCCNGCNGCARGCN 1020

Db 1059 ccaactgcagagccctctggaaccacatctgtgcaggtggccaccactgaagtcataagc 1118
Qy 1021 CCNACNGCARGCNGWSNWSNCCNACNWSNCTNCARGTNCNCCNACNACNGARGTNAHWSN 1080

Db 1119 actgcaactgtgcagatgcccaactgcagagacacaggtatgacacctgagaagtgcca 1178
Qy 1081 ACNGCNCNGTNCARATGCCNACNGCNGARGWSNACNGNATGACNCCNGARAARGTNCNCN 1140

Db 1179 gtttcagagtggtgggtaccacactggcagagatgtcaactccagaggtcacaggtatg 1238
Qy 1141 GTNWSNARGTNATGGGNCNACNCTYTNCGNGARATGHSNACNCCNCGARGCNGCNGNATG 1200

Db 1239 acactgcagaggtatcaattgtggtcttctggaaccacacagctgcacaggtacaact 1298
Qy 1201 ACNCCNGCARGTNSNATHGCTGNTYTNWSNGNACNACNGCNGCNGCARGTNAACNACN 1260

Db 1299 acagtggtgggtggagaccacagctagagagctacctatccctgacctgaaggtccagat 1358
Qy 1261 ACNGARTGGGTNGARACNACNCGNMGNGARYTNCNATHCCNARGCNGARGCNGCNGAY 1320

Db 1359 gccagctcaatcatgtctacggaaagtattacaggttccctggggcccccctgctggtggt 1418
Qy 1321 GGNWSNWSNATHGWSNACNARGWSNATHACNGGWSNYTNGCNGCNYTNTYTNAYGNGN 1380

Db 1419 acagccaccttaagtggtggaagacaaagtccccctggattgtgtctgtatcgatat 1478
Qy 1381 ACNCGNACNCTYTMGNYTNGTNAARMGNCARGTNCNCCNCTNCAYTGTNTYTNAYMGTAY 1440

Db 1479 ggttcttttccgtcacctggacattgtccaggggtattgaaagtgcgagatccctcgag 1538
Qy 1441 GGNWSNTTYSNCTNACNYTNGAYATHGTCNARGGNATHGARMWSNCGNGARATHYTNCAR 1500

Db 1539 gctgtccgtccggtgagggggatgcatcttgagctgactgtccttgcacaggcgggctg 1598
Qy 1501 CGCTGTCNWSNWSNCGARGGNGAYCGCTTYGARYTNACNCTNWSNTGYCARGCNGCNGYTN 1560

Db 1599 cccaaggaagcctgcattggagatctcatcgccaggtggccagccccctgcccagcggtg 1658
Qy 1561 CCNAAARGCNGTGYATGGARATHWSNWSNCCNGGNTGYCARCCNCCNGCNGCARMGNYTN 1620

Qy 661 CARYTMNGNCTYNGAYGGNGNAAVACARCAITTYTNNMNAAYCARCONYNACNTY 720

Db 742 gccctccagctccatgaccccaaggctatctggctgaagctgaacctctctacactgg 801

Qy 721 GONTYTCARYTNCAYGAYCWNWSNGNTAYTNGCNGARGCNGAYTNNWSNTAYACNTG 780

Db 802 gacttggagacagtagtgaaccctgatctctcggcgaacttggctgaactcactaac 861

Qy 781 GAYTTYGGNGAYWSNWSNGNACNYTNAHWSMNGCNYTNGTNGTNAACNAYACNTAY 840

Db 862 ctggagctggcccaagctcaactcccagctgctcctgagctgccattctctcactcc 921

Qy 841 YTNACRCNGCNGCNGTNAACNACNACNACNACNACNACNACNACNACNACNACN 900

Db 922 tctggctctcccaagctccagcaccacagatggcacagagcccaactcgagagccct 981

Qy 901 TCGGNWSNWSNCCNGTNCNGNACNACNACNACNACNACNACNACNACNACNACN 960

Db 982 aacaccacagctggcccaagtgcctactacagaagttgtgggtactaacactggtcagcg 1041

Qy 961 AAYACNACNGCNGCNGCARGTNCNACNACNACNACNACNACNACNACNACNACN 1020

Db 1042 ccaactgcagagccctctggaaccacatctgtgcaggtgccaccactgaagtcataagc 1101

Qy 1021 CONACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 1080

Db 1102 actgcaactgtgcagatgccaaactgcagagcagcaggtatgacacctgagaagtgcca 1161

Qy 1081 ACNCCNGCNGTNCARATGCCNACNACNACNACNACNACNACNACNACNACNACN 1140

Db 1162 gttccagaggtcatgggtaccacactggcagagatgtcaactccagagctacaggtatg 1221

Qy 1141 GTNWSNARGTNAATGGNACNACNACNACNACNACNACNACNACNACNACNACN 1200

Db 1222 acaactgcagaggtatcaattgtgtgtcttctggaaccacagctgcacaggtacaact 1281

Qy 1201 ACNCCNGCNGARGTNSNATHCTGTNTYTNWSNACNACNACNACNACNACNACN 1260

Db 1282 acagagtggtggagaccacagctagagactacatccctgagctgaaggtccagat 1341

Qy 1261 ACNARGTGGTNGARACNACNACNACNACNACNACNACNACNACNACNACNACN 1320

Db 1342 gccagctcaatcatgtctacggaagtattacaggttccctggcccccctgctggatggt 1401

Qy 1321 GGNWSNWSNATHATGWSNACNACNACNACNACNACNACNACNACNACNACNACN 1380

Db 1402 acagccacactaaagctggtgaagacaaagtcccccctggattgtgtctgatcatgat 1461

Qy 1381 ACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 1440

Db 1462 gttctcttccctcaacctggacattgtccaggtattgaaagtgcagagatccctgcag 1521

Qy 1441 GGNWSNTYWSNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 1500

Db 1522 gctgctgcctggaggaggtgacttggagctgactgtgctgcgaagcgggctg 1581

Qy 1501 GCGTNCNWSNNGCNGARGCNGAYCCTYTGARYTNACNCTNWSNTGYCARGCNGNYN 1560

Db 1582 cccaggaagctgcatggagatctcatccaggggtgccagccccctggccagcggctg 1641

Qy 1561 CNAARGCNGNTGATGARGATHWSNWSNCCNGTGCARCCNCCNGCNGCARGMNTN 1620

Db 1642 tgcagctgtgctaccagccagctgcagctggttctgcacagactgaaggtg 1701

Qy 1621 TCGARCCGNTYTNCCNWSNCCNGTGCARYTNGTNTNCAVCARATHYTNARGN 1680

Db 1702 ggcctgggacatactgctcaatggtctctggctgataccaacagctggcagtggtc 1761

Qy 1681 GGNWSNACNACNTAYTGYTNAAYTNWSNNTYNGCNGAYACNAAYTWSNNTYNGCNGTGN 1740

Db 1762 agcaccagcttatactgctgctcaagaagcagccttggcagagttccgctgactgctg 1821

Qy 1741 WSNACNARYTNATHATGCCNGCNGARGCNGGNYTNGCNGCARGTNCNCCNYTNATHGN 1800

Db 1822 ggcatactgctggtggtgctgctgctgctgctgctgctgctgctgctgctgctg 1881

Qy 1801 GGNATHYTNNGTNTYTNATGCCNCTGNTYTNNGCNSNYTNATHATYHGMNMGNYN 1860

Db 1882 atgaacgaagcttctccagcttaccagcttgcacacagcagcagctgctgctgctg 1941

Qy 1861 ATGAARCAAGATTTYSNCTNCCNARYTNCCNCAYSNWSNWSNCAITGGYTNMGNTN 1920

Db 1942 cccgcacatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2001

Qy 1921 CCNMGNATHYTYGWSNTGYCNAHGCNGARAYWSNCCNYTNTNWSNCCNARGAR 1980

Db 2002 gt 2003

Qy 1981 CT 1982

RESULT 3

ID T03760 standard; DNA; 2131 BP.

AC T03760;

DT 25-MAR-1996 (first entry)

DE Melanoma-specific immunogen, pMEL17.

KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;

KW cytotoxic T cell; lymphocyte; HLA-A2; ss.

OS Homo sapiens.

PN W09522561-A2.

PD 24-AUG-1995.

PF 16-FEB-1995; U01991.

PR 16-FEB-1994; US-197399.

PR 29-APR-1994; US-234784.

PA (UYVI-) UNIV VIRGINIA PATENT FOUND.

PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;

DR WPI; 95-302688/39.

PT Melanoma-specific immunogen comprises epitope(s) homologous with

PT pMEL17 - are highly potent stimulators of HLA-A2+CTL's useful in

PT adoptive immuno-therapy

PS Disclosure; Page 19-20; 148pp; English.

CC A melanoma-specific immunogen homologous with pMEL-17 (t03760) comprises

CC one or more CTL (cytotoxic T lymphocyte) epitopes from the group

CC R82098-R82194 capable of eliciting a CTL response. The epitopes R82098-

CC R82108 are of particular interest. The immunogen can be used for

CC partial protection in mammals against melanoma peptides which are

CC homologous with pMEL-17 are highly potent stimulators of HLA-A2+

CC CTLs in several cell lines and can be used in immunotherapy or

CC incorporated into immunogenic conjugates as vaccines.

SQ Sequence 2131 BP; 474 A; 589 C; 577 G; 491 T;

Query Match 87.9%; Score 6699; DB 17; Length 2131;

Best Local Similarity 56.2%; Pred. No. 0.00e+00;

Matches 1125; Conservative 416; Mismatches 440; Indels 21; Gaps 1;

Db 13 tggatctggtgctaaaaagatgctcttcttctgctgctgctgctgctgctgctgctg 72

Qy 2 TCGAYTNGTNTNARGNNTGYTNTNCAITNGCNGTNGCNGCNYTNTYNGCNG 61

	Query Match	2.1%;	Score 157;	DB 17;	Length 1065;	
	Best Local Similarity	9.6%;	Pred. No. 7.46e-01;			
	Matches 21;	Conservative 61;	Mismatches 137;	Indels 0;	Gaps 0;	
D b	70	ttyathmgntngdngcngwsgargarytngntttggyacnattaytgacnytngar	129			
	: : :	: : : :	: : : :	: : : :	:	:
C <p></p>	1744	TWMSNACNCGCNARNSRRTTCTGCGNARNWSNACRTTNARRCARTANGTCNCNW	1685			
D b	130	caycayctyacngarttggynytnacngnaaytnttgytngcngcngaaytyn	189			
	: : :	: : : : :	: : : : :	: : : : :	:	:
C <p></p>	1694	SNCNCYTTTARDATYTGTGCNARNACNARYTGRCANGCNGNSGNRNACNGGYT	1625			
D b	190	gmnmgnacnaarcyntncargtngnacnatggndgtngtyccnagncncaycn	249			
	: : :	: : :	: : :	: :	: :	:
C <p></p>	1624	GRCARNRCKYTGNGCNGGGYGTGRCCNCGNWSNSDATYTCCATRCANGCYTYT	1565			
D b	250	gtmngncarytnrgarygtntytnytngaycaratg	288			
	: : :	: : : :	: : : :	: : : :	:	:
C <p></p>	1564	TNGENARNCNCYTCGCCANWSNACNGNTNARYTCRAAG	1526			

	7	RESULT
ID	T05466 standard; cDNA to mRNA; 36 BP.	
AC	T05466;	
DT	25-JAN-1996 (first entry)	
DE	Sequence encoding immunogenic peptide of melanoma antigen gp100.	
DE	Melanoma; antigen; vaccine; immunogen; primer; probe; detection;	
KW	kw identification; tumour; gp100; es.	

Accession	Gene	Protein	Location/Qualifiers
FT	Key		
FT	CDS		1..36
FT	/*tag= a		
FT	/product=	Immunogenic peptide.	
FT	protein_bind		1..33

FT	/tag= b	
FT	protein_bind	7..36
FT	/tag= c	
FT	protein_bind	7..33
FT	/tag= d	
FT	protein_bind	10..36
FT	/tag= e	
PN	EP-668350-A1.	
PD	23-AUG-1995.	
PF	14-FEB-1995; 200348.	
PR	16-FEB-1994; EP-200337.	
PR	21-DEC-1994; EP-203709.	
PA	(ALKU) AKZO NOBEL NV.	
PI	Adema GJ, Figdor CG;	
DR	WPI; 95-284790/38.	
DR	P-PSDB; R78642.	
PT	Melanoma associated antigen gp100 - used in vaccines and for the	
PS	detection of tumours	
PS	Claim 7; Page 27; 40pp; English.	
CC	Immunogenic peptides derived from the melanoma associated antigen	
CC	(See R78639-45) may be used in the production of vaccines.	
CC	Nucleotide sequences encoding the immunogenic peptides may be used	
CC	as primers and probes in the detection of melanoma cells. Tumour	
CC	infiltrating lymphocytes capable of binding to the melanoma	
CC	associated antigen can be cultured ex vivo and returned to melanoma	
CC	particles, and when radiolabelled, they may be used to identify	
CC	tumour deposits.	
SQ	Sequence 36 BP; 9 A; 8 C; 11 G; 8 T;	

Query Match 2.0%; Score 150; DB 16; Length 36;

ID	Q03969	standard; DNA; 909 BP.
AC	Q03969;	
CT	AC	
DT	22-AUG-1989	(first entry)
DE	Clone P25a	
EW	Enrich ascites tumour; protein p25; ss.	
KW	Mus musculus.	
FH	Key	Location/Qualifiers
FT	CS	29.,751
FT	CS	/*tag= a
FT	/label=P25	protein
PN	DD-273071-A.	
PD	01-NOV-1989.	

Key	Location/Qualifiers
FT Peptide	1..22
/label= signal sequence	
FT Protein	12
/label= mature NPBR	
FT Domain	23..455
/label= extracellular domain	
FT /note= "binds natriuretic peptides A,B and C"]	
FT Domain	456..456
/label= transmembrane domain	
FT Domain	479..1047
/label= cytoplasmic domain	
FT /note= "GC and protein kinase activity"	
FT Modified -site	24..26
/label= N-glycos site	
FT Modified -site	35..37
/label= N-glycos site	
FT Modified -site	161..163
/label= N-glycos site	
FT Modified -site	195..197
/label= N-glycos site	
FT Modified -site	244..246
/label= N-glycos site	
FT Modified -site	277..279
/label= N-glycos site	
FT Modified -site	349..351
/label= N-glycos site	

```
FT /label= N-glycos_site
FT Modified-site 600..602
FT /label= N-glycos_site
PN W09100292-A.
PD 10-JAN-1991.
PF 22-JUN-1990; U03586.
PR 23-JUN-1989; US-370673.
PA (GETH ) GENENTECH INC.
PI Chang M, Goeddel D, Lowe D;
DR WP1; 91-036711/05.
DR N-PSDB; Q10324.
PT Natriuretic protein receptor B - for diagnosis and treatment of
PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS Claim 3; Fig 1; 49pp; English.
CC The sequence was derived from the DNA encoding natriuretic peptide
CC receptor B, NPRB, having guanylyl cyclase (GC) activity and protein
CC kinase activity. The DNA can be inserted into expression vectors
CC for the prodn. of the protein, opt. after being mutated to produce
CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC 114,952). The protein (or variants) can be used in treatment of
CC natriuretic peptide disorders, and also to isolate peptides using
CC affinity chromatography. Antibodies with affinity for NPRB can
CC also be prepd.
SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;

Query Match 1.9%; Score 144; DB 2; Length 1047;
Best Local Similarity 15.5%; Pred. No. 5.00e+00;
Matches 52; Conservative 104; Mismatches 178; Indels 2; Gaps 2;

Db 345 anvnntnnngtndgrnrvnkmgrryhgvtgvnmvdkndndrntdvnwngdndsgd 404
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Cp 1982 ACYTGTCGNCNWSNARNRNWNSRRTTYTTCNCCDATNGRCANWSRCARAADATCKN 1923

Db 405 mnaahysganknwrfgrnnnwfwgkannndnncandnddhsdcktmstnanvgtg 464
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Cp 1922 GGNARNGKNA-RCCARTGNWSNWSRGTGNGNARTYTGNGGNACNWSRAARTCYGT 1864

Db 465 ntunmgvsnnnrnknmnknnaemwrnrwnnnnngsnryhkgagstntnsgesy 524
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Cp 1863 CATNARNCKNCKRTADATNARNWSNGCNARNACNCGCCATNARNACNARNARDAT 1804

Db 525 genmtahqkynnnaantghnkgvkvankhvkknrrntrnnnnkhmdvnnhnrtrng 584
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Cp 1803 NCCNACDATNARNNGNACYTGNCNARNCCNGCYTCYTGNCNCGCATDATNARYTNGT 1744

Db 585 acndnnncvntnrcrgsnndnnnnndndndmwnrysnndndvkgmannhnnenshgs 644
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Cp 1743 NWSNACNACNAGNARNWSRTGTTCGCGNARNWSNACRTTNARNRCARTANGTCNWS 1684

Db 645 nksnscvndsrnvknkd-ygnasrstamddna 679
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Cp 1683 NCCNCCYTNNARDATYTGRTGNARNACNARYTGCA 1648
```

```
RESULT 11
ID NR1094 standard; DNA; 202 BP.
AC NR1094;
DT 29-OCT-1990 (first entry)
DE Sequence of a synthetic crystalline 'silk' gene based on protein fibroid
DE of silk fibre produced by Bombyx mori moth larvae
KW Structural protein; microbial silk production; ds.
OS Bombyx mori.
FH Key Location/Qualifiers
FT misc_feature 1..4
```

```
FT /tag= a
FT /note="Sticky end"
FT misc_feature complement (199..202)
FT /tag= b
FT /note="Sticky end"
FT misc_feature 1..23
FT /tag= c
FT /label=Oligo PA9
FT misc_feature 24..46
FT /tag= d
FT /label=Oligo PA11
FT misc_feature 47..74
FT /tag= e
FT /label=Oligo PA13
FT misc_feature 75..99
FT /tag= f
FT /label=Oligo PA15
FT misc_feature 100..128
FT /tag= g
FT /label=Oligo PA17
FT misc_feature 129..150
FT /tag= h
FT /label=Oligo PA19
FT misc_feature 151..176
FT /tag= i
FT /label=Oligo PA21
FT misc_feature 177..202
FT /tag= j
FT /label=Oligo PA23
FT misc_feature complement (5..27)
FT /tag= k
FT /label=Oligo PA10
FT misc_feature complement (28..50)
FT /tag= l
FT /label=Oligo PA12
FT misc_feature complement (51..78)
FT /tag= m
FT /label=Oligo PA14
FT misc_feature complement (79..103)
FT /tag= n
FT /label=Oligo PA16
FT misc_feature complement (104..132)
FT /tag= o
FT /label=Oligo PA18
FT misc_feature complement (133..154)
FT /tag= p
FT /label=Oligo PA20
FT misc_feature complement (155..180)
FT /tag= q
FT /label=Oligo PA22
FT misc_feature complement (181..202)
FT /tag= r
FT /label=Oligo PA24
PN EP-294979-A.
PD 14-DEC-1988.
PF 31-MAY-1988; 304906.
PR 10-JUN-1987; GB-013559.
PA (PACO-) PA Consult Services.
PI Edwards RM, Light JA, Nicholson K;
DR WP1; 88-355406/50.
PT Recombinant plasmids contg. DNA region and lac Z gene portion -
PT produce novel silk-like structural proteins similar to Bombyx
PT mori fibroid
PS Claim 7; Page 5; 17pp; English.
```

ID Q59596 standard; cDNA; 405 BP.
AC Q59596;

OS	Key	Location/Qualifiers
FS	uninsure	17..19
FT	/*tag= a	
FT	/note= "encodes Leu"	
FT	uninsure	26..28
FT	/*tag= b	
FT	/note= "encodes Arg"	
FT	uninsure	29..31
FT	/*tag= c	
FT	/note= "encodes Leu"	
FT	uninsure	32..34
FT	/*tag= d	
FT	/note= "encodes Leu"	
FT	uninsure	38..40
FT	/*tag= e	
FT	/note= "encodes Leu"	
FT	uninsure	41..43

Aug 27 18:59

US-08-231-565A-27.rng

23

```

PR 12-SEP-1985; US-774359.
PR 30-MAY-1986; US-868473.
PR 11-JUL-1986; US-884460.
PR 04-JUN-1987; US-054991.
PR 04-MAY-1988; US-190293.
PR 08-FEB-1991; US-654397.
PR 25-SEP-1991; US-765472.
PR 25-SEP-1992; US-951365.
PA (MER1 ) MERCK & CO INC.
PI Gimenez-Callego G, Kelly LJ, Linemeyer DL, Thomas KA;
DR WPI; 95-138983/18.
DR P-PSDB; R74648.
PT New recombinant human acidic fibroblast growth factor - used to
PT promote cell growth, to promote wound healing, for vascular
PT grafts and blood vessel repair
PS Disclosure; Table 1; 25pp; English.
CC Oligonucleotides were synthesized on the basis of the amino
CC acid sequence of cattle acidic fibroblast growth factor (aFGF)
CC and used to construct a synthetic gene (given in Q88233) that
CC incorporates codons preferred by E. coli or mammalian cells,
CC unique cloning sites, etc. The preferred sequence is given in
CC Q88234. The synthetic gene can be mutated to obtain a human aFGF
CC synthetic gene for production of recombinant protein (R74647) .
SQ Sequence 420 BP; 87 A; 37 C; 62 G; 55 T;

Query Match 1.7%; Score 128; DB 14; Length 420;
Best Local Similarity 6.8%; Pred. No. 4.65e+01;
Matches 12; Conservative 63; Mismatches 102; Indels 0; Gaps 0;

Db 178 garacngncarttytngcncatggaacnagvggnyntntayggwnancarcncn 237
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 1966 RNANGGWSRTTYTNCDDATNGRCANWSRCARADATNCKNGNARCNCKNCCART 1907
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 238 aaygargartgytnttytgarmgnytgargaraaycaytayaacntayathwn 297
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 1906 GNWSNWSNWSRTGNGGNARYTGNGGNACNWSRAARTCYTTCATNARNCKNCKNCKRT 1847
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 298 aaraarcaycngaraarcaytggttytgngnytnaaraaraayggmgnwna 354
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 1846 ADATNARNWSGCCNARNACNACGCCATNARNACNARNARDATNCCNACDATNARN 1790

```

Search completed: Tue Aug 27 19:11:30 1996
Job time : 172 secs.

WATERMAN

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_tpn n.a. - n.a. Smith-Waterman search, using a protein query
which has been backtranslated into n.a. using IUPAC symbols

Run on: Tue Aug 27 19:11:52 1996; MaePar time 767.73 Seconds
Tabular output not generated. 917.837 Million cell updates/sec

Title: >US-08-231-565A-27

Description: (1-661) from US08231565A.pep

Perfect Score: 7622

N.A. Sequence: 1 ATGCAAYTNGTNTNARMG.....TNTYNSNGNCARCARGTN 1983

Comp: TACCTRRACNANRNTTKK.....ANRANSWNCNGTGTTCAN

Scoring table: TABLE bktranslate2

Gap 30

Mmatch STD : Dbase 0; Query 0

Searched: 509049 seqs, 177673129 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:STS1 92:STS2
93:STS3 94:STS4 95:STS5 96:STS6 97:STS7
EST-STS-TWO
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100
108:EST101 109:gnEST1 110:gnEST2 111:gnEST3 112:gnEST4
113:gnEST5 114:gnEST6 115:gnEST7 116:gnEST8 117:gnEST9
118:gnEST10 119:gnEST11 120:gnEST12 121:gnEST13
122:gnEST14 123:gnEST15 124:gnSTS1 125:gnSTS2 126:gnEST1

Database:

EST-STS-TWO
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100
108:EST101 109:gnEST1 110:gnEST2 111:gnEST3 112:gnEST4
113:gnEST5 114:gnEST6 115:gnEST7 116:gnEST8 117:gnEST9
118:gnEST10 119:gnEST11 120:gnEST12 121:gnEST13
122:gnEST14 123:gnEST15 124:gnSTS1 125:gnSTS2 126:gnEST1

RESULT 1

127:enEST2 128:enEST3 129:enEST4 130:enEST5 131:enEST6
132:enEST7 133:enEST8 134:enEST9 135:enEST10 136:enEST11

Statistics: Mean 69.416; Variance 71.291; scale 0.974

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	1939	25.4	621	114	N28728	y*67a12.r1 Homo sapie	0.00e+00
2	1939	25.4	621	136	HS728262	y*67a12.r1 Homo sapie	0.00e+00
3	1812	23.8	638	114	N28736	y*67c12.r1 Homo sapie	0.00e+00
4	1812	23.8	638	136	HS736264	y*67c12.r1 Homo sapie	0.00e+00
5	1510	19.8	546	136	HS746262	y*67f11.r1 Homo sapie	0.00e+00
6	1510	19.8	546	114	N28746	y*67f11.r1 Homo sapie	0.00e+00
7	1350	17.7	521	118	N40290	y*80c07.r1 Homo sapie	0.00e+00
8	1350	17.7	521	131	HS290283	y*80c07.r1 Homo sapie	0.00e+00
9	1297	17.0	416	68	R78477	y173d06.r1 Homo sapie	0.00e+00
10	1006	13.2	306	83	T29726	EST92123 Homo sapiens	0.00e+00
11	974	12.8	286	46	R01509	ye79a02.r1 Homo sapie	3.85e-302
12	948	12.4	430	113	N23118	y*67c12.s1 Homo sapie	6.28e-292
13	948	12.4	430	130	HS118284	y*67c12.s1 Homo sapie	6.28e-292
14	877	11.5	431	113	N23111	y*67a12.s1 Homo sapie	3.75e-264
15	877	11.5	431	129	HS111256	y*67a12.s1 Homo sapie	3.75e-264
16	292	3.8	506	130	HS123264	y*47e01.r1 Homo sapie	1.70e-48
17	292	3.8	506	114	N29123	y*47e01.r1 Homo sapie	1.70e-48
18	173	2.3	495	19	HS1731	yp85g02.r1 Homo sapie	3.27e-13
19	171	2.2	58	130	HS128259	y*67f11.s1 Homo sapie	1.07e-12
20	171	2.2	58	113	N23128	y*67f11.s1 Homo sapie	1.07e-12
21	154	2.0	522	81	T22039	4047 Arabidopsis thal	1.66e-08
22	147	1.9	336	84	T34583	EST71046 Homo sapiens	7.10e-07
23	144	1.9	493	113	N24647	y*89g12.s1 Homo sapie	3.40e-06
24	144	1.9	493	135	HS647262	y*89g12.s1 Homo sapie	3.40e-06
25	138	1.8	127	5	H03700	yj42f09.r1 Homo sapie	7.15e-05
26	140	1.8	288	83	T31913	EST40905 Homo sapiens	2.62e-05
27	138	1.8	306	93	G05045	human STS WI-6432	7.15e-05
28	137	1.8	338	40	HSDEK123	H. sapiens partial cD	1.17e-04
29	134	1.8	354	35	HSC17H061	H. sapiens partial cD	5.10e-04
30	141	1.8	360	121	CEIK087C2F	C.elegans cDNA clone	1.58e-05
31	141	1.8	360	127	CEK087C2F	C.elegans cDNA clone	1.58e-05
32	134	1.8	364	108	T99127	ye62d08.r1 Homo sapie	5.10e-04
33	139	1.8	389	6	H10164	ym05e09.s1 Homo sapie	4.34e-05
34	139	1.8	442	55	R36012	yg69a03.r1 Homo sapie	4.34e-05
35	134	1.8	450	87	T44050	7313 Arabidopsis thal	5.10e-05
36	138	1.8	495	116	N34300	yy51f12.s1 Homo sapie	7.15e-05
37	138	1.8	495	131	HS300270	yy51f12.s1 Homo sapie	7.15e-05
38	136	1.8	498	47	R04666	pk27b11.s1 Caenorhabd	1.92e-04
39	138	1.8	519	9	H18327	yn49a07.s1 Homo sapie	7.15e-05
40	133	1.7	192	12	H26106	y144f12.s1 Homo sapie	8.25e-04
41	133	1.7	371	108	T97699	ye54c03.s1 Homo sapie	8.25e-04
42	133	1.7	383	93	G07853	human STS CHLC.GATA72	8.25e-04
43	132	1.7	431	27	H78737	yu49f07.r1 Homo sapie	1.33e-03
44	132	1.7	511	98	T62611	yc04c12.r1 Homo sapie	1.33e-03
45	133	1.7	543	8	HS5883	ym22e11.r1 Homo sapie	8.25e-04

ALIGNMENTS

Db
61 aaagtattacaggttccttgggccccctgctggatggtacaggccaccttaaggctggtga 120
||:: ||:|| :| :| || :| ||:|| ||:|| ||:|| :| :| :| :|

FH	Key	Location/Qualifiers
001		

Query match 15.2%; score 1000; DB 63; length 300;
Best Local Similarity 54.4%; Pred. No. 0.00e+00;

IMAGE Consortium (info@image.llnl.gov) for further information.

23
24

=> d his

(FILE 'USPAT' ENTERED AT 10:53:45 ON 27 AUG 96)

L1 155 S MART OR GP100
L2 7055 S NUCLEIC ACID
L3 4343 S CDNA
L4 8911 S L2 OR L3
L5 14 S L1 AND L4
L6 7 S L1(L)L4
L7 0 S L1(P)L4
L8 7 S L6

=> d 1-7 cit

1. ~~5,484,891~~, Jan. 16, 1996, Selectin ligands; Laurence A. Lasky, et al., 530/387.3; 435/7.2; 530/350, 395 [IMAGE AVAILABLE]

2. ~~5,464,815~~, Nov. 7, 1995, Inhibition of heparin-binding; Steven Chamow, et al., 514/8; 424/85.2; 436/86, 87; 514/21; 530/412 [IMAGE AVAILABLE]

3. ~~5,369,025~~, Nov. 29, 1994, Recombinant fowlpox vaccine for protection against Marek's disease; Keyvan Nazerian, et al., 435/235.1; 424/186.1, 199.1, 229.1; 435/69.3, 172.3, 240.2; 530/350; 536/23.72; 935/9, 32, 34, 57, 63, 65, 70 [IMAGE AVAILABLE]

4. ~~5,318,890~~, Jun. 7, 1994, Assays for inhibitors of leukocyte adhesion; Steven Rosen, et al., 435/7.24, 7.1, 7.2, 7.92; 530/387.3 [IMAGE AVAILABLE]

5. ~~5,304,640~~, Apr. 19, 1994, DNA sequence encoding a selectin ligand; Laurence A. Lasky, et al., 536/23.5; 435/69.1, 172.3, 240.2, 320.1 [IMAGE AVAILABLE]

6. ~~5,242,829~~, Sep. 7, 1993, Recombinant pseudorabies virus; Dennis L. Panicali, et al., 435/320.1; 424/199.1, 229.1, 232.1; 435/69.1, 69.3, 172.3 [IMAGE AVAILABLE]

7. ~~5,141,922~~, Aug. 25, 1992, Biologically active proteins and a method for their use; Gwen G. Krivi, 514/12, 21; 530/324, 399 [IMAGE AVAILABLE]

=>

=> d his

(FILE 'USPAT' ENTERED AT 10:46:09 ON 27 AUG 96)

E KAWAKAMI, Y/IN

E ROSENBERG, S/IN

L1 24 S E15,E14
L2 0 S MELANONA(3A)ANTIGEN#(5A)(LYMPHOCYTE# OR LYMPHOKYTE#)
L3 3 S MELANOMA(3A)ANTIGEN#(5A)(LYMPHOCYTE# OR LYMPHOKYTE#)
L4 7055 S NUCLEIC ACID
L5 4343 S CDNA
L6 8911 S L4 OR L5
L7 0 S L6 AND L3
L8 9 S L1 AND L6
L9 0 S L3 AND L8

=>